التوزيعات الإحصائية

تطبيقات وتقديرات المعالم

تأليف

Nick T. Thomopoulos

ترجمة

د. منصور بن محمد شراحيلي

دار جامعة الملك سعود للنشر

KING SAUD UNIVERSITY PRESS

التوزيعات الإحصائية تطبيقات وتقديرات المعالم

تأليف Nick T. Thomopoulos

ترجمة د. منصور بن محمد شراحيلي

2020

بسم الله الرحمن الرحيم

التوزيعات الإحصائية

تطبيقات وتقديرات المعالم

تأليف

Nick T. Thomopoulos

ترجمة

د. منصور بن محمد شراحيلي

أستاذ مساعد بقسم الإحصاء وبحوث العمليات

كلية العلوم - جامعة الملك سعود

دار جامعة الملك سعود للنشر

KING SAUD UNIVERSITY PRESS

ص.ب ٦٨٩٥٣ - الرياض ١١٥٣٧ المملكة العربية السعودية

فهرسة مكتبة الملك فهد الوطنية أثناء النشر

ثوموبولوس، نيك ت.

التوزيعات الإحصائية: تطبيقات وتقديرات المعالم / نيك ت. ثوموبولوس؛ منصور بن محمد شراحيلي - الرياض، ١٤٤١هـ.

٢١٥ ص؛ ١٧سم × ٢٤ سم

ردمك: ٧ - ٨٣٣ - ٥٠٧ - ٦٠٣ - ٩٧٨

١ - الإحصاء الرياضي أ. شراحيلي، منصور بن محمد (مترجم) ب. العنوان

ديوي ٥١٩ ١٤٤١/٤٦٤٦

رقم الإيداع: ١٤٤١/٤٦٤٦

ردمك: ٧ - ٨٣٣ - ٥٠٧ - ٦٠٣ - ٩٧٨

هذه ترجمة عربية محكمة صادرة عن مركز الترجمة بالجامعة لكتاب:

Statistical Distributions: Applications and Parameter Estimates
By: Nick T. Thomopoulos.
© Springer International Publishing AG 2017.

وقد وافق المجلس العلمي على نشرها في اجتماعه الثاني للعام الدراسي ١٤٤٠/ ١٤٤١هـ، المعقود بتاريخ ١٧/ ١/ ١٤٤١هـ، الموافق ١٦/ ٩/ ٢٠١٩م.

دار جامعة الملك سعود للنشر
KING SAUD UNIVERSITY PRESS

مقدمة المترجم

الحمد لله رب العالمين والصلاة والسلام على خاتم النبيين والمرسلين سيدنا محمد عليه أفضل الصلاة وأتم التسليم وبعد، لا يخفى على القارئ العزيز أهمية التوزيعات الاحتمالية سواءً كانت متقطعة أم متصلة وما لها من تطبيقات عديدة في شتى مجالات الحياة والدور الكبير الذي تقدمه في عملية اتخاذ القرار حول ظاهرة معينة، لذا نقدم للمكتبة العربية هذا العمل والذي يحتوي على أهم التوزيعات الاحتمالية المستخدمة في الحياة، ونبذة عن كل توزيع من خلال تقدير معالمه حيث إنه في معظم الدراسات التطبيقية لا تكون هذه المعالم معروفة وذلك عن طريق إيجاد أهم المقاييس العددية الخاصة به من خلال استخدام بيانات عينة من هذا التوزيع، كما يشمل بعض التطبيقات بالإضافة إلى عرض الشكل البياني الخاص بكل توزيع عند قيم محددة للمعالم.

وفي الختام أتقدم بالشكر الجزيل لكل من ساهم بتحسين هذا العمل وأبدى ملاحظاته لتطويره سائلاً المولى عز وجل أن يكون هذا العمل إضافة جيدة للمكتبة العربية وأن ينفع به المهتمين في التحليل الإحصائي والاحتمالي إنه جواد كريم.

المترجم

إهداء المؤلف

لزوجتي، وأولادي، وأحفادي.

شكر وتقدير

أتوجّه بالشكر على وجه الخصوص لزوجتي، ايلين ثوموبولوس، التي شجعتني على كتابة هذا الكتاب، والتي قدمت لي المشورة كلّما دعت الحاجة، وأيضاً أتوجه بالشكر لدانيال سوسمان على تدقيقه للنص. كما أشكر العديد من الأشخاص الذين ساعدوني وكانوا مصدر إلهامٍ بالنسبة لي على مرّ السنين، بمن فيهم بعض طلاب الدكتوراه في معهد إلينوي للتكنولوجيا. ولا يمكنني إلا ذكر بعض الأسماء هنا: إيمانويل بيتينيس (الجامعة الوطنية للعلوم الصحية)، وفريد بوك (معهد أبحاث إلينوي للتكنولوجيا)، وديك تشيابيتا (شيابيتا وويلش)، وآل إندرس (جامعة تامبا)، وجون جاروفالاكيس (جامعة باتراس)، وجيمس هول (كايوود شيلر أسوشيتس)، ومونتيرا جانتارافاريرات (معهد إلينوي للتكنولوجيا)، وأرفيد جونسون (جامعة سان فرانسيس)، وكارول ليندي (بانديت)، وأناتول لونجينو (معهد إلينوي للتكنولوجيا)، وفوتس موزاكيس (فراين ريسيرتش)، وجورج ريسنيكوف (جامعة كاليفورنيا)، وبول سبيراكيس (جامعة باتراس).

نبذة عن المؤلف

حصل نيك ت. ثوموبولوس على درجة البكالوريوس في الأعمال ودرجة الماجستير في الرياضيات من جامعة إلينوي، ودرجة الدكتوراه في الهندسة الصناعية من معهد إلينوي للتكنولوجيا. كان مشرفاً على بحوث العمليات في معهد هارفستر العالمي، وكبير العلماء في معهد أبحاث إلينوي للتكنولوجيا، وبروفسور في الهندسة الصناعية، في كلية ستيوارت للأعمال في معهد إلينوي للتكنولوجيا. قام بتأليف أحد عشر كتاباً بما في ذلك: أساسيات نظم صفوف الانتظار (سبرنجر)، وأساسيات محاكاة مونت كارلو (سبرنجر)؛ وطرائق التنبؤ التطبيقي (برنتايس هول)، وأساسيات الإنتاج والمخزون، وسلسلة التوريد (أتلانتيك). كما قام بنشر العديد من الأبحاث، وقدّم استشارات عديدة في مجموعة واسعة من القطاعات في الولايات المتحدة الأمريكية، وأوروبا، وآسيا. كما حصل نيك على العديد من الجوائز على مرّ السنين، مثل: جائزة ريست من جمعية أبحاث العمليات العسكرية للتطورات الجديدة التي أدخلها على نظرية صفوف الانتظار، وجائزة الأستاذ المميّز في بانكوك، تايلاند، من جمعية الخريجين الآسيويين من معهد إلينوي للتكنولوجيا، وجائزة الإنجاز المهني من جمعية الخريجين الآسيويين من معهد إلينوي للتكنولوجيا.

مقدمة المؤلف

التوزيع الإحصائي هو دالّة رياضية تعرّف حدوث محتمل لمتغيّر عشوائي عبر مداه المقبول. ويعد فهم التوزيعات الإحصائية مطلباً أساسياً للباحثين في جميع التخصصات تقريباً حيث سيختار الباحث المطّلع التوزيع الإحصائي الذي يناسب البيانات الموجودة في الدراسة. يقدّم هذا الكتاب وصفاً لمجموعة التوزيعات الإحصائية ذات التطبيق الواسع في الدراسات التي تبحث في الإحصاءات والاحتمالات. إنَّ بعض التوزيعات معروفة للباحث العام ويتم استخدامها بعدة طرق، في حين أنَّ التوزيعات المفيدة الأخرى يصعب فهمها وليست شائعة الاستخدام. يُبيّن هذا الكتاب متى وكيف يمكن تطبيق كلّ نوع من التوزيعات في الدراسات البحثية بهدف تحديد التوزيع الأفضل تطبيقاً على الدراسة. إنَّ التوزيعات الإحصائية هي للمتغيرات العشوائية المتصلة، والمنفصلة وثنائية المتغيرات. وفي معظم الدراسات، تكون قيم المعالم غير معروفة مسبقًا، وهناك حاجة إلى بيانات العيّنة لتقدير قيم هذه المعالم. وفي سيناريوهات أخرى لا تتوفر بيانات العيّنة، ويسعى الباحث لبعض المعلومات التي تساعد على تقدير قيم المعالم. إنَّ هذا الكتاب سهل القراءة ويتضمن العديد من الأمثلة التي ترشد القارئ وسيكون مرجعاً مفيداً للغاية لأي شخص يقوم بإجراء تحليل إحصائي واحتمالي. ويشمل ذلك علماء الإدارة، والباحثين في السوق، والمهندسين، وعلماء الرياضيات، والفيزيائيين، والكيميائيين، والاقتصاديين، والباحثين في العلوم الاجتماعية، والطلاب في العديد من التخصصات.

نيك ت. ثوموبولوس

بير ريدج، إلينوي، الولايات المتحدة الأمريكية

المحتويات

الفصل الأول

المفاهيم الإحصائية
STATISTICAL CONCEPTS

١-١ مقدمة

التوزيع الإحصائي هو دالّة رياضية تحدّد كيفية حدوث نتائج محاولة تجريبية عشوائياً بطريقةٍ محتملة. تُسمى هذه النتائج بالمتغيرات العشوائية وتقع منطقة التعريف في فضاء عينة محدد يرتبط كل منها بتوزيع فردي. وغالباً ما تنقسم التوزيعات الإحصائية إلى نوعين: متصلة ومنفصلة. تُطبّق التوزيعات الاحتمالية المتصلة عندما يقع المتغير العشوائي بين حدين مثل كمية مياه الأمطار التي تتجمع في وعاء بسعة 5 غالون بعد هطول المطر. بينما يُطبّق التوزيع الاحتمالي المنفصل عندما تكون نتائج التجربة قيماً محددة، مثل عدد النقاط التي تظهر على حجري نرد. كما يمكن تصنيف التوزيعات على أنها وحيدة المتغير أو متعددة المتغيرات، بحيث يكون التوزيع وحيد المتغير عندما يكون للتوزيع متغير واحد فقط ويكون التوزيع متعدد المتغيرات عندما يكون للتوزيع متغيران عشوائيان أو أكثر. في هذا الكتاب سوف نتطرق إلى التوزيعات الاحتمالية المتصلة والمنفصلة أحادية المتغير شائعة الاستخدام، وإلى التوزيعات الإحصائية المتصلة ثنائية المتغير التي تطبّق بصورةٍ متكررة، والتي لها متغيران عشوائيان مرتبطان بصورةٍ مشتركة.

١-١-١ التوزيعات الاحتمالية، والمتغيرات العشوائية، والترميز، والمعالم

تم إدراج التوزيعات الاحتمالية ومعالم هذه التوزيعات والتعبير عنها فيما يلي:

التوزيعات المتصلة:

التوزيع المنتظم المتصل	$X \sim CU(a,b)$
التوزيع الأسي	$X \sim Exp(\theta)$
توزيع إرلانج	$X \sim Erl(k,\theta)$
توزيع جاما	$X \sim Gam(k,\theta)$
توزيع بيتا	$X \sim Beta(k_1,k_2,a,b)$
توزيع ويبل	$X \sim We(k_1,k_2,\gamma)$
التوزيع الطبيعي	$X \sim N(\mu,\sigma^2)$
توزيع اللوغاريتم الطبيعي	$X \sim LN(\mu_y,\sigma_y^2)$
التوزيع الطبيعي المقطوع الأيسر	$T \sim LTN(k)$
التوزيع الطبيعي المقطوع الأيمن	$T \sim RTN(k)$
التوزيع المثلثي	$X \sim TR(a,b,\tilde{x})$

التوزيعات المنفصلة:

التوزيع المنتظم المنفصل	$X \sim DU(a,b)$
توزيع ذي الحدين	$X \sim Bin(n,p)$
التوزيع الهندسي	$X \sim Ge(p)$
توزيع باسكال	$X \sim Pa(k,p)$
التوزيع الهندسي الزائدي	$X \sim HG(n,N,D)$
توزيع بواسون	$X \sim Po(\theta)$

التوزيعات ثنائية المتغير:

التوزيع الطبيعي ثنائي المتغير	$X_1,X_2 \sim BVN(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$
توزيع اللوغاريتم الطبيعي ثنائي المتغير	$X_1,X_2 \sim BVLN(\mu_{y1},\mu_{y2},\sigma_{y1},\sigma_{y2},\rho_y)$

ونوردُ فيما يلي التوزيعات المتصلة، مع عرض موجزٍ عن كل نوع:

التوزيع المنتظم المتصل	تكون الكثافة أفقية.
التوزيع الأسي	تبلغ الكثافة ذروتها عند الصفر وتنحدر بعد ذلك.

توزيع إرلانج	تتراوح العديد من الأشكال من أسي إلى طبيعي.
توزيع جاما	تتراوح العديد من الأشكال من أسي إلى طبيعي.
توزيع بيتا	تنحرف العديد من الأشكال نحو اليسار، واليمين، وتكون متماثلة.
توزيع ويبل	تتراوح العديد من الأشكال من أسي إلى طبيعي.
التوزيع الطبيعي	يكون على شكل جرس متماثل.
توزيع اللوغاريتم الطبيعي	يبلغ ذروته قرب الصفر وينحرف إلى أقصى اليمين.
التوزيع الطبيعي المقطوع الأيسر	يكون التوزيع الطبيعي مقطوعاً جهة اليسار وينحرف نحو اليمين.
التوزيع الطبيعي المقطوع الأيمن	يكون التوزيع الطبيعي مقطوعاً جهة اليمين وينحرف نحو اليسار.
التوزيع المثلثي	ترتفع الكثافة نحو الذروة، ثم تنحدر نحو الصفر.

كما نوردُ فيما يلي التوزيعات المنفصلة، مع موجزٍ عن كل نوع:

التوزيع المنتظم المنفصل	تكون الاحتمالية أفقية.
توزيع ثنائي الحدين	محاولات n مع احتمال ثابت للنجاح لكل محاولة.
التوزيع الهندسي	عدد المحاولات حتى أول نجاح.
توزيع باسكال	عدد المحاولات حتى نجاح k.
توزيع بواسون	عدد الأحداث عندما يكون معدل الحدث ثابتاً.
التوزيع الهندسي الزائدي	عيّنات n دون إرجاع الكثير من حجم N.

وأخيراً نوردُ فيما يلي التوزيعات ثنائية المتغير، مع عرض موجزٍ عن كل نوع:

التوزيع الطبيعي ثنائي المتغير	يتم تشكيل التوزيعات الهامشية بشكلٍ طبيعي.
توزيع اللوغاريتم الطبيعي ثنائي المتغير	تكون التوزيعات الهامشية لوغاريتم طبيعي.

١-٢ أساسيات

يصف هذا الفصل الخصائص المتعلقة بالتوزيعات الإحصائية المتصلة والمنفصلة التي تُستخدم عادةً، ومن هذه الخصائص الدوال الاحتمالية، والمتوسط، والتباين، والانحراف المعياري، والمنوال و الوسيط. وعندما تتوفر بيانات العيّنة، فإنها تُستخدم لمساعدة المحلّل على تقدير قيم معالم التوزيع الإحصائي قيد الدراسة. ويتمثلَ تقدير عيّنة من القياسات في إيجاد القيمة الدنيا، والقيمة العليا، والمتوسط، والتباين، والانحراف المعياري، والمنوال والوسيط.

١-٣ التوزيع المتصل

المدى المقبول: يكون للتوزيع المتصل متغير عشوائي X، ويكون المدى المقبول له على النحو التالي:

$$a \leq x \leq b$$

حيث قد يكون a عدداً لا نهائياً سالباً وقد يكون b عدداً لا نهائياً موجباً.

كثافة الاحتمال: إن دالة الكثافة الاحتمالية لـلمتغير العشوائي X هي:

$$f\ (x) \geq 0, \quad a \leq x \leq b$$

حيث إن

$$\int_a^b f(x)dx = 1$$

التوزيع التراكمي: إن دالة التوزيع التراكمية للمتغير العشوائي X هي:

$$F(x) = \int_a^x f(w)dw$$

إنَّ ذلك يجعل الاحتمال التراكمي للمتغير X الأقل من x_0 أو المساوي لها، على النحو أدناه:

$$F(x_0) = P(X \leq x_0) = \int_a^{x_0} f(x)dx$$

الاحتمال المكمل: يتم الحصول على دالة الاحتمال المكمل عندما تكون X أكبر من x_0 على النحو التالي:

$$H(x_0) = 1 - F(x_0) = P(X > x_0)$$

القيمة المتوقعة: يتم اشتقاق $E(X)$ وهي القيمة المتوقعة لـلمتغير العشوائي X، كما يرمز لمتوسط X بالرمز μ ويعرف على النحو التالي:

$$\mu = E(X) = \int_a^b x f(x) dx$$

التباين والانحراف المعياري: يتم الحصول على تباين X ، أو σ^2 على النحو التالي:

$$\sigma^2 = E\left[\left(X-\mu\right)^2\right] = \int_a^b \left(x-\mu\right)^2 f(x) dx$$

ومنها فإن الانحراف المعياري σ يساوي

$$\sigma = \sqrt{\sigma^2}$$

الوسيط: إن وسيط المتغير العشوائي X والذي يُشار إليه بالرمز $\mu_{0.5}$، هو قيمة X مع الاحتمال التراكمي 0.50 كما هو مبيّن أدناه:

$$F(\mu_{0.5}) = 0.5$$

المنوال: يعرف المنوال $\tilde{\mu}$ ، على أنه القيمة الأكثر احتمالاً للمتغير X ، ويوجد حيث تكون دالة الكثافة الاحتمالية في أقصاها في المدى المقبول كما هو مبيّن أدناه:

$$f(\tilde{\mu}) = max\left\{f(x),\ a \le x \le b\right\}$$

يتم الحصول على النقطة المئوية α للمتغير X ، والتي يُشار إليها $x\alpha$ ، من خلال الدالة المعكوسة $F(x)$ حيث إن

$$F(\alpha x) = \alpha$$

ملاحظة: على سبيل المثال وسيط المتغير X هو $X_{0.5} = \mu_{0.5}$

معامل الاختلاف: إنَّ معامل اختلاف المتغير العشوائي X هو نسبة الانحراف المعياري إلى المتوسط كما هو مبيّن أدناه:

$$COV = \frac{\sigma}{\mu}$$

١-٤ التوزيعات المنفصلة

المدى المقبول: إنَّ للتوزيع المنفصل متغيراً عشوائياً X ، ذا مدى مقبول

$$a \leq x \leq b$$

لتبسيط الأمر، يقتصر المدى في هذا الكتاب على الاختلاف بزيادات واحدة، أي أن المدى يساوي

$$\left(a, a+1, \ldots\ldots, b-1, b\right)$$

الدالة الاحتمالية: تكون دالة الكتلة الاحتمالية $P(x)$ للمتغير X ، كما هو مبيّن أدناه:

$$P(x) \geq 0, \quad a \leq x \leq b$$

حيث إنَّ

$$\sum_{a}^{b} P(x) = 1$$

الاحتمال التراكمي: إنَّ دالة الاحتمال التراكمية $F(x)$ لـلمتغير العشوائي X ، تعرف على النحو التالي:

$$F(x) = \sum_{a}^{x} P(w)$$

وهو ما يعطي الاحتمال التراكمي للمتغير $X = x_0$ أو أقل، كما هو مبيّن أدناه:

$$F(x_0) = P\left(X \leq x_0\right) = \sum_{a}^{x_0} P(x)$$

الاحتمال المكمّل: إنَّ دالة الاحتمال المكمّل للقيمة x_0 ، هي احتمال أن يكون X أكبر من x_0 كما هو مبيّن أدناه:

$$H(x_0) = 1 - F(x_0) = P(X > x_0)$$

القيمة المتوقعة والمتوسط: يتم اشتقاق القيمة المتوقعة $E(X)$ للمتغير X، وتُسمى أيضاً متوسط X، μ، على النحو المبيّن أدناه:

$$\mu = E(X) = \sum_{a}^{b} xP(x)$$

التباين والانحراف المعياري: يتم الحصول على التباين σ^2 المتغير العشوائي X على النحو التالي:

$$\sigma^2 = E\left[\left(X-\mu\right)^2\right] = \sum_{a}^{b}\left(x-\mu\right)^2 P(x)$$

و من ثم يتم حساب الانحراف المعياري σ كما هو مبيّن أدناه:

$$\sigma = \sqrt{\sigma^2}$$

الوسيط: إنَّ وسيط المتغير العشوائي X، الذي يُشار إليه بواسطة $\mu_{0.5}$، هو قيمة X مع الاحتمال التراكمي 0.50 كما هو مبيّن أدناه:

$$F(\mu_{0.5}) = 0.5$$

المنوال: إنَّ المنوال $\tilde{\mu}$ هو القيمة الأكثر احتمالاً للمتغير X ويوجد حيث يكون الاحتمال أقصاه ضمن المدى المقبول كما هو مبيّن أدناه:

$$P(\tilde{\mu}) = max\left\{P(x),\ a \le x \le b\right\}$$

نسبة لكسيس (Lexis Ratio): إنَّ نسبة لكسيس τ بالنسبة للمتغير X هي نسبة اختلاف التباين إلى المتوسط كما هو مبيّن أدناه:

$$\tau = \frac{\sigma^2}{\mu}$$

١-٥ الإحصاءات الأساسية لبيانات العيّنة

عند جمع بيانات العيّنة n، $\left(x_1, \ldots, x_n\right)$، يمكن حساب مختلف القياسات الإحصائية كما هو مبيّن أدناه:

n = حجم العيّنة

$x(1)$ = القيمة الدنيا للقيم $(x_1, \ldots, x_n)$

$x(n)$ = القيمة العليا للقيم $(x_1, \ldots, x_n)$

متوسط العيّنة = $\overline{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$

تباين العيّنة = $S^2 = \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2 \ / \ n - 1$

الانحراف المعياري للعيّنة = $S = \sqrt{S^2}$

معامل الاختلاف = $\text{cov} = \dfrac{S}{\overline{x}}$

منوال العيّنة = $\tilde{x}$

وسيط العيّنة = $x_{0.5}$

وأخيراً نسبة لكسيس = $\tau = \dfrac{S^2}{\overline{x}}$

إنَّ وسيط العيّنة هو القيمة الوسطى لمجموعة البيانات المصنّفة والمرتبة تصاعدياً أو تنازلياً $(x_1, \ldots, x_n)$. والتي يتم إدراجها على النحو التالي: $x(1), x(2), \ldots, x(n)$. إذا كان عدد العيّنات فردياً، فإنَّ وسيط العيّنة هو:

$$x_{0.5} = x(n`)$$

حيث $n` = \dfrac{(n+1)}{2}$

وإذا كان عدد المشاهدات n زوجياً، فإنَّ الوسيط هو:

$$x_{0.5} = \frac{\left[x(n \ / \ 2) + x(n + 1 \ / \ 2)\right]}{2}$$

إنَّ منوال العيّنة $\tilde{x}$ هو قيمة البيانات الأكثر تكراراً من بيانات العيّنة. أحياناً قد يظهر لدينا منوالان أو أكثر، وأحياناً لا يظهر أي منوال. لإيجاد المنوال ينبغي على المحلل تصنيف البيانات واختيار القيمة التي تظهر بأكثر تكرار. وإذا لم تظهر أي قيمة أكثر من غيرها من القيم،

يمكن تصنيف البيانات إلى مـجموعات ومن ثم اختيار متوسط المجموعـة ذات التكرار الأكبر علـى أنه المنوال.

١-٦ طرق تقدير المعالم

عندما يريد المحلل تطبيق توزيع احتمالي ما في دراسة بحثية غالباً ما تكون قيمة (قيم) المعلمة (المعالم) مجهولة وينبغي تقديرها. يتم إيجاد التقدير بصورةٍ عامة من بيانات العيّنة $(x_1, \ldots, x_n)$ التي تم جمعها. إنَّ الطريقتين المعروفتين لتقدير المعالم من مدخلات البيانات هما طريقة مقدّر الإمكان الأعظم وطريقة العزوم. نُوردُ فيما يلي وصفاً مختصراً لكلٍّ منهما.

١-٦-١ طريقة مقدّر الإمكان الأعظم (MLE)

تصيغ هذه الطريقة دالة احتمالية باستخدام بيانات العيّنة $(x_1, \ldots, x_n)$ ومعلمة (معلمات) التوزيع قيد الدراسة، وتسعى للحصول على قيمة المعلمة (المعالم) التي تعظّم هذه الدالة. على سبيل المثال، عندما يكون للتوزيع الإحصائي معلمة واحدة θ، يكون البحث عن قيمة θ التي تجعل هذه الدالة أكبر ما يمكن باستخدام n من العينات، وتُسمى هذه القيمة مقدّر الإمكان الأعظم.

١-٦-٢ طريقة العزوم (MoM)

تستخدم هذه الطريقة لإيجاد العزوم النظرية للتوزيع من دالة الاحتمال، ويتم إيجاد عزوم العيّنة المقابلة من بيانات العيّنة $(x_1, \ldots, x_n)$، باستخدام دالة التوزيع الاحتمالي. تعطي مجموعة العزوم النظرية معالم المجتمع (θ, σ, μ، وغير ذلك). وينتج عن استبدال عزوم العيّنة بالعزوم النظرية تقدير هذه المعالم ($\hat{\theta}, S, \tilde{x}$ وغير ذلك)، ويُسمى هذا التقدير بطريقة العزوم.

١-٧ تحويل المتغيرات

من الضروري في بعض الأحيان تغيير بيانات العيّنة الأصلية $(x_1, \ldots, x_n)$ بطريقة تتيح تحديد أسهل لتوزيع الاحتمال الذي يلائم بيانات العيّنة بشكلٍ أفضل. هناك طريقتان مفيدتان في هذه الحالة و هما، التحويل إلى مجموعة بيانات الصفر أو أكبر، والتحويل إلى مجموعة الصفر وواحد. نورد فيما يلي وصفاً لهاتين الطريقتين.

١-٧-١ تحويل البيانات إلى صفر أو أكبر

قد يجد المحلل من خلال بيانات العيّنة $(x_1, \ldots, x_n)$ أنه من المفيد تحويل البيانات إلى مجموعة جديدة حيث إنَّ جميع المدخلات هي صفر أو أكبر. وذلك من خلال تطبيق:

$$y = x - x(1)$$

على كل مُدخل، وحيث إن $x(1)$ هي القيمة الصغرى لمجموعة بيانات x. تصبح المجموعة الجديدة لمدخلات البيانات n : $(y_1, \ldots, y_n)$. ويصبح متوسط مجموعة البيانات الجديدة y والانحراف المعياري لها على النحو التالي:

$$\overline{y} = \overline{x} - x(1)$$

$$S_y = S_x$$

أحياناً يكون هناك حاجة إلى إيجاد معامل الاختلاف من مجموعة بيانات y لتحديد توزيع الاحتمال المناسب لبيانات العيّنة. ويصبح معامل الاختلاف لمجموعة البيانات الجديدة y على النحو التالي:

$$COV_y = \frac{S_y}{\overline{y}}$$

١-٧-٢ تحويل البيانات إلى صفر وواحد

أحياناً يكون من المفيد تحويل بيانات العيّنة إلى مجموعة جديدة توجد ضمن مدى صفر إلى واحد. يتم تحقيق ذلك من خلال تطبيق ما يلي:

$$W = \left[x - x(1)\right] / \left[x(n) - x(1)\right]$$

على كل قيمة في مجموعة بيانات العيّنة، حيث إن $x(1)$ هي القيمة الصغرى، و $x(n)$ هي القيمة القصوى. وتصبح مجموعة البيانات الجديدة على النحو التالي: $(w_1, \ldots, w_n)$. ومن هذه الطريقة يكون متوسط w والانحراف المعياري له على النحو التالي:

$$\overline{w} = \left[\overline{x} - x(1)\right] / \left[x(n) - x(1)\right]$$

$$S_w = S_x \;/\; \left[x(n) - x(1)\right]$$

كما يصبح معامل الاختلاف من مجموعة البيانات w على النحو التالي:

$$COV_w = \frac{S_w}{\overline{w}}$$

من خلال مجموعة البيانات w الموجودة ضمن (1،0)، يكون معامل الاختلاف الذي يظهر مفيداً أحياناً لتحديد توزيع احتمال مجموعة البيانات.

المثال (١-١). باعتبار بيانات العيّنة ذات الأحد عشر مدخلاً والمدرجة على النحو التالي: $(x_1, \ldots, x_{11}) = (14, 23, 26, 31, 27, 22, 15, 17, 31, 29, 34)$. يتم إدراج القياسات الإحصائية الأساسية للبيانات كما هو مبيّن أدناه:

$$n = 11$$

$$x(1) = 14$$

$$x(11) = 34$$

$$\overline{x} = 24.45$$

$$S_x^2 = 46.87$$

$$S_x = 6.85$$

$$COV_x = 0.28$$

$$\tilde{x} = 31$$

$$x_{0.5} = 26$$

المثال (١-٢). على افتراض أن محلل البيانات في المثال (١-١) يريد تحويل البيانات لإنتاج مجموعة جديدة حيث إن $y \geq 0$. يتم تحقيق ذلك من خلال أخذ القيمة الدنيا $x(1) = 14$، وتطبيق $y = x - 14$ على كل مدخل فنحصل على المجموعة المؤلفة من 11 مدخلاً ذات المتغير X والمجموعة المقابلة ذات المتغير Y أدناه:

$$(x_1, \ldots, x_{11}) = (14, 23, 26, 31, 27, 22, 15, 17, 31, 29, 34)$$

$$(y_1,\ldots,y_{11}) = (0,9,12,17,13,8,1,3,17,15,20)$$

تم إدراج القياسات الإحصائية الأساسية لمجموعة البيانات المحولة y أدناه. ويلاحظ أنَّ القياسات الوحيدة للمجموعة y التي لم تزل ثابتة مقارنة مع بيانات X هي عدد العناصر، والتباين، والانحراف المعياري.

$$n = 11$$
$$y(1) = 0$$
$$y(11) = 20$$
$$\overline{y} = 10.45$$
$$S_y^2 = 46.87$$
$$S_y = 6.85$$
$$COV_y = 0.65$$
$$\tilde{y} = 17$$
$$y_{0.5} = 12$$

المثال (١-٣). على افتراض أن محلل البيانات في المثال (١-١) يريد تحويل البيانات الموجودة بين 0 و1. يتم تحقيق ذلك من خلال أخذ القيمة الدنيا = $x(1)$=14 والقيمة العليا = $x(11)$=34 وتطبيق العلاقة (14– 34) / (14– x) =w على كل مدخل، ونورد فيما يلي القيم الأصلية لـ x ومجموعة w المحوّلة.

$$(x_1,\ldots,x_{11}) = (14,23,26,31,27,22,15,17,31,29,34)$$

$$(w_1,\ldots,w_{11}) = (0.45,0.00,0.60,0.85,0.65,0.40,0.05,0.15,0.85,0.75,1.00)$$

وتكون القياسات الإحصائية الأساسية على النحو التالي:

$$n = 11$$
$$w(1) = 0$$

$$w(11) = 1$$

$$\overline{w} = 0.52$$

$$S_w^2 = 0.34$$

$$S_w = 0.65$$

$$COV_w = 1.25$$

$$\tilde{w} = 0.85$$

$$w_{0.5} = 0.60$$

١-٧-٣ التوزيعات المتصلة ومعامل الاختلاف

عندما يكون لدى المحلل بيانات العيّنة ويسعى للحصول على التوزيع الإحصائي المناسب لتطبيقه على البيانات، يكون معامل الاختلاف مفيداً أحياناً في تحديد التوزيعات المتصلة التي قد تناسب بيانات العينات. نورد فيما يلي بعض التوزيعات المرشّحة للقيم المختارة من معامل الاختلاف.

عندما يكون معامل الاختلاف ≤ 1.00: التوزيع الأسي، وتوزيع إرلانج ($k_1 = 1$)، وتوزيع جاما ($k_1 \leq 1$)؛ وتوزيع ويبل ($k_1 \leq 1$)، وتوزيع اللوغاريتم الطبيعي، والتوزيع الطبيعي المقطوع الأيسر، والتوزيع الطبيعي المقطوع الأيمن، وتوزيع بيتا.

عندما يكون معامل الاختلاف ≥ 0.33: التوزيع الطبيعي، وتوزيع إرلانج ($k_1 = 9$)، وتوزيع جاما ($k_1 \geq 9$)، وتوزيع ويبل ($k_1 \geq 4$)، وتوزيع بيتا.

عندما يكون معامل الاختلاف بين (0.33 و1.00): التوزيع المنتظم المتصل، وتوزيع إرلانج، وتوزيع جاما، وتوزيع بيتا، والتوزيع الطبيعي المقطوع الأيسر، والتوزيع الطبيعي المقطوع الأيمن ، وتوزيع ويبل.

١-٧-٤ التوزيعات المنفصلة ونسبة لكسيس

عندما يكون لدى الباحث بيانات عيّنات منفصلة ويسعى للحصول على توزيعات منفصلة تناسب هذه البيانات، تكون نسبة لكسيس مفيدة أحياناً في تحديد التوزيعات المرشحة للاستخدام. نورد فيما يلي قائمة بقيم نسبة لكسيس والتوزيعات المرشحة:

عندما يكون $\tau < 1$: توزيع ثنائي الحدين.

عندما يكون $\tau = 1$: توزيع بواسون.

عندما يكون $\tau \geq 1$: التوزيع الهندسي مع عدد الإخفاقات على أنه المتغير، وتوزيع باسكال مع عدد الإخفاقات على أنه المتغير.

١-٨ ملخص

إنَّ لتوزيعات الاحتمال أحادية المتغير متغيراً واحداً وهي ذات نوعين، متصلة و منفصلة. في حين أن للتوزيعات الطبيعية ثنائية المتغير وتوزيعات اللوغاريتم الطبيعي ثنائية المتغير اثنين من المتغيرات ذات الصلة بشكل مشترك. كما أنَّ للتوزيعات المتصلة دالة احتمال تحدد كيفية الحصول على النتائج الممكنة بطريقة الاحتمال. كذلك للتوزيعات المنفصلة دالة الكتلة الاحتمالية التي تحدد احتمال كل نتيجة خاصة في مدى العيّنة. وتكون القياسات الإحصائية المشتركة هي المتوسط، والتباين، والانحراف المعياري، والوسيط، والمنوال، ومعامل الاختلاف، ونسبة لكسيس. وعندما تكون قيم المعالم مجهولة وبيانات العيّنة متاحة، فإنه يتم حساب تقدير القياسات الإحصائية.

الفصل الثاني

التوزيع المنتظم المتصل
CONTINUOUS UNIFORM DISTRIBUTION

يستخدم التوزيع المنتظم المتصل عندما يقع المتغير العشوائي في أي مكان بين حدين، وغالباً ما يستخدم التوزيع عندما لا يكون لدى المحلل معلومات قطعية عن مدى وشكل المتغير العشوائي. على سبيل المثال، قد تقدّر الإدارة الزمن اللازم لإنهاء المشروع على الأرجح ما بين 50 و60 ساعة، أو على سبيل المثال يُقدر المسؤولون المسافة التي قطعتها كرة السلة وحطّمت الرقم القياسي بين 410 و430 قدماً، أيضاً كمية الثلوج المتساقطة في موقع ما في أحد الأيام الشتوية تتراوح بين 1 و5 إنش. يحتوي هذا الفصل على دالة الكثافة الاحتمالية، والتوزيع التراكمي، والمتوسط، والتباين، والانحراف المعياري للمتغير العشوائي. كما يصف النقطة المئوية α لـلمتغير X التي تحدد قيمة x حيث إنّ الاحتمال التراكمي لها يساوي α . كذلك عندما تكون قيم المعالم مجهولة وبيانات العينة متاحة، يتم الحصول على تقدير هذه المعالم حيث يتم استخدام إحدى الطريقتين في التقدير من خلال طريقة مقدّر الإمكان الأعظم، والأخرى من خلال طريقة العزوم، وعند عدم توفر بيانات العيّنة، يتم أخذ رأي الخبراء لوضع التقدير المناسب. غالباً ما يكون كلا الحدين مجهولين ويكون هناك حاجة لكلا التقديرين. أحياناً تكون القيمة الدنيا فقط معلومة، وفي بعض الأحيان تكون القيمة العليا فقط مجهولة. لذا تم وصف طريقة تقدير قيم المعلمتين بالنسبة للسيناريوهات الثلاثة.

٢-١ أساسيات

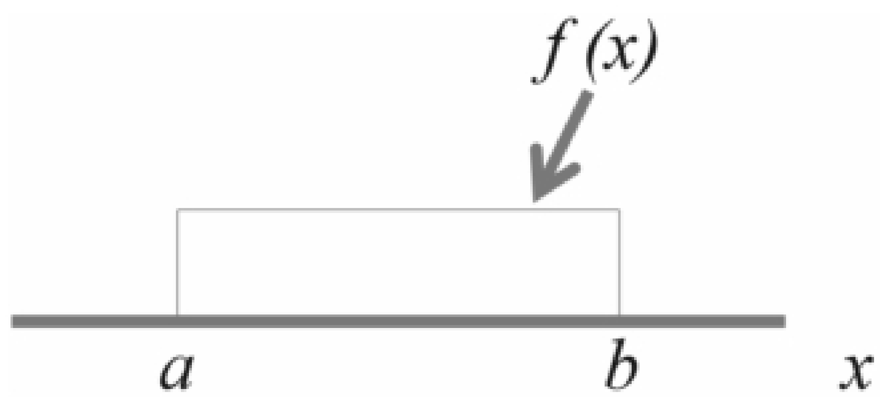
إنَّ المتغير العشوائي X الذي يتبع للتوزيع المنتظم المتصل له مدى مقبول (من a إلى b)، حيث إنه من المرجح أن تحدث أي قيمة في هذا المدى. كما تحدد معالم التوزيع نطاق المدى وهي:

a = القيمة الدنيا

b = القيمة العليا

نورد فيما يلي دالة الكثافة الاحتمالية، وتمثيلاً لها في الشكل (٢-١).

$$f(x) = 1 \,/\, (b-a), \quad a \le x \le b$$



شكل (٢-١). دالة الكثافة المنتظمة المتصلة.

كما نورد فيما يلي المتوسط، والتباين، والانحراف المعياري للتوزيع المنتظم المتصل:

$$\mu = (a+b) \,/\, 2$$

$$\sigma^2 = (b-a)^2 \,/\, 12$$

$$\sigma = (b-a) \,/\, \sqrt{12}$$

ويتم الحصول على معامل الاختلاف كما هو مبيّن أدناه:

معامل الاختلاف = $\dfrac{\sigma}{\mu}$

في الحالة الخاصة عندما يكون $a=0$ و$b=1$، يصبح معامل الاختلاف مساوياً لـــما يلي

معامل الاختلاف = $2 \,/\, \sqrt{12} = 0.578$

وتكون دالة التوزيع التراكمي $F(x)$ معرفة على النحو التالي:

$$F(x) = (x-a) \,/\, (b-a) \;, \quad a \le x \le b$$

إنَّ النقطة المئوية α على x، والمُشار إليها بواسطة $x\alpha$، مرتبطة بالاحتمال التراكمي α، كما هو مبيّن أدناه:

$$P(X \leq x\alpha) = \alpha$$

ولإيجاد قيمة $x\alpha$ بالنسبة للتوزيع المنتظم المتصل، نقوم بتطبيق ما يلي:

$$x\alpha = a + \alpha(b - a)$$

المثال (٢-١). في نظام الإنتاج، تختلف كمية السائل المسكوب في أحد الأوعية وقد يكون من 7.0 إلى 7.2 أونصة. وبالنسبة لوعاء تم أخذه عشوائياً كعيّنة، نجد أن كمية السائل في هذا الوعاء والمشار إليها بالمتغير X، المدى المقبول: $7 \leq X \leq 7.2$، وتكون دالة الكثافة الاحتمالية على الصورة:

$$f(x) = 1 \,/\, (7.2 - 7) = 5.0 \quad 7.0 \leq X \leq 7.2$$

إن متوسط كمية السائل في هذا الوعاء هو $\mu = (7.0 + 7.2) \,/\, 2 = 7.10$. ويصبح التباين $\sigma^2 = (7.2 - 7.0)^2 \,/\, 12 = 0.0033$ والانحراف المعياري $\sigma = \sqrt{0.0033} = 0.0577$.

كما يصبح التوزيع الاحتمالي التراكمي لأي قيمة من قيم x في المدى المقبول بالصورة التالية:

$$F(x) = (x - 7.0) \,/\, (0.20) \qquad 7.0 \leq x \leq 7.2$$

نلاحظ أنه على سبيل المثال، يتم الحصول على احتمال X الأدنى أو المساوي لـ 7.15 كما هو مبيّن أدناه:

$$P(X \leq 7.15) = F(7.15) = (7.15 - 7) \,/\, 0.20 = 0.75$$

على سبيل المثال، يتم حساب قيمة النقطة 0.25% بالنسبة للمتغير X على النحو التالي:

$$x_{0.25} = 7.0 + 0.25(7.20 - 7.00) = 7.05$$

٢-٢ بيانات العيّنة

عندما يريد المحلل تطبيق التوزيع المنتظم المتصل في إحدى الدراسات بحيث كانت قيم المعلمتين (b، a) مجهولة، عندئذ يتم استخدام بيانات العيّنة لتقدير هذه المعالم. وهي المشاهدات العشوائية n والمُشار إليها بواسطة $(x_1, \ldots,\ x_n)$، وللحصول على التقدير، تؤخذ الإحصائيات التالية من بيانات العيّنة:

$\overline{x}$ = المتوسط

S = الانحراف المعياري

$x(1)$ = القيمة الدنيا

$x(n)$ = القيمة العليا

المثال (٢-٢). ينتج عن تجربة عشوائية ما مشاهدات $n = 10$ كما يلي:

$$9.1, 3.1, 17.1.15.8, 12.6.5.9, 5.1, 14.2, 19.8, 7.3$$

يفترض المحلل أن يتم الحصول على البيانات من التوزيع المنتظم المتصل ويريد إيجاد منتصف مدى القيم 50%. تتطلب الخطوة الأولى تقدير المعلمتين $(a،b)$، ولتحقيق ذلك فإنَّ المهمة الأولى هي قياس الإحصائيات الأساسية من البيانات كما يلي:

$$\overline{x} = 11.00$$

$$S = 5.69$$

$$x(1) = 3.1$$

$$x(n) = 19.8$$

٢-٣ تقدير المعالم من بيانات العيّنة

هناك طريقتان لتقدير المعالم. الطريقة الأولى هي طريقة مقدّر الإمكان الأعظم، والثانية هي طريقة العزوم.

يتم تقدير المعالم باستخدام طريقة مقدّر الإمكان الأعظم على النحو التالي:

$$\hat{a} = x(1)$$

$$\hat{b} = x(n)$$

كما يصبح تقدير المعالم باستخدام طريقة العزوم كما هو مبيّن أدناه:

$$\hat{a} = \overline{x} - \sqrt{12}S \ / \ 2$$

$$\hat{b} = \overline{x} + \sqrt{12}S \ / \ 2$$

المثال (٢-٣). في المثال السابق (٢-٢)، يكون التقدير من خلال طريقة مقدّر الإمكان الأعظم ومن خلال طريقة العزوم، كما هو مبيّن أدناه:

باستخدام طريقة مقدّر الإمكان الأعظم، تصبح مقدرات المعالم كالتالي:

$$\hat{a} = x(1) = 3.1$$

$$\hat{b} = x(n) = 19.8$$

يتم حساب منتصف المدى 50٪ على النحو التالي:

$$x_{0.25} \leq x \leq x_{0.75}$$

حيث إن $x_{0.25}$ و x_{075} هما قيمتا النقطتين 0.25% و0.75% بالنسبة للمتغير X، على التوالي. باستخدام طريقة مقدّر الإمكان الأعظم، تصبح القيمتين على النحو التالي:

$$x_{0.25} = \hat{a} + 0.25(\hat{b} - \hat{a}) = 3.1 + 0.25(19.8 - 3.1) = 7.27$$

$$x_{0.75} = \hat{a} + 0.75(\hat{b} - \hat{a}) = 3.1 + 0.75(19.8 - 3.1) = 15.62$$

وبالتالي يصبح منتصف المدى 50% كما يلي

$$7.27 \leq x \leq 15.62$$

ويكون تقدير الاحتمال المُرتبط ضمن المدى على النحو التالي:

$$P(7.27 \leq x \leq 15.62) = 0.50$$

باستخدام البيانات من المثال (٢-٢)، يتم حساب تقدير المعالم باستخدام طريقة العزوم على النحو التالي:

$$\hat{a} = \overline{x} - \sqrt{12}S \ / \ 2 = 11.00 - \sqrt{12} \times 5.69 \ / \ 2 = 1.14$$

$$\hat{b} = \overline{x} + \sqrt{12}S \ / \ 2 = 11.00 + \sqrt{12} \times 5.69 \ / \ 2 = 20.86$$

وبتطبيق تقدير طريقة العزوم، يصبح منتصف المدى 50% على النحو التالي:

$$x_{0.25} = \hat{a} + 0.25\left(\hat{b} - \hat{a}\right) = 1.14 + 0.25(20.86 - 1.14) = 6.07$$

$$x_{0.75} = \hat{a} + 0.75\left(\hat{b} - \hat{a}\right) = 1.14 + 0.75(20.86 - 1.14) = 15.69$$

٢-٤ تقدير المعالم عند عدم توفر بيانات

باعتبار الحالة التي يريد فيها المحلل تطبيق التوزيع المنتظم المتصل ولكن لا يوجد لديه تقدير للمعلمتين $(a،b)$ وليس لديه بيانات العينات لأخذ التقدير. عندئذ يطلب المحلل مشورة الخبراء الذين يقدمون بعض التقديرات ضمن مدى المتغير X .

٢-٥ عندما تكون المعلمتان (a,b) مجهولتين

عندما تكون قيم المعلمتين مجهولتين، يطلب المحلل مشورة الخبير الذي يعطي تقديراً للنقط المئوية للمتغير X ، والمشار إليها بالرمز (x_1 = النقطة المئوية α_1 ، و x_2 = النقطة المئوية α_2). نلاحظ أن:

$$P(X < x_1) = \alpha_1$$

$$P(X < x_2) = \alpha_2$$

إذا كان $\alpha_1 = 0$ ، فإن تقدير القيمة الدنيا هو $\hat{a} = x_1$ ؛ وإذا كان $\alpha_2 = 1$ ، فإن تقدير القيمة العليا هو $\hat{b} = x_2$.

نعرض فيما يلي كيفية تقدير قيم المعالم الدنيا والعليا عندما يكون $\alpha_1 \geq 0$ ، $\alpha_2 \leq 1$ و $\alpha_2 > \alpha_1$. باستخدام $x_1, x_2, \alpha_1, \alpha_2$ ، يتم الحصول على تقدير المعالم كما هو مبين أدناه مع ملاحظة أن x_1 و x_2 مرتبطان بالمعلمتين (b ، a):

$$x_1 = a + \alpha_1(b - a)$$

$$x_2 = a + \alpha_2(b - a)$$

نلاحظ أيضاً العلاقات المكافئة أدناه:

$$x_1 = a(1-\alpha_1) + \alpha_1 b$$

$$x_2 = a(1-\alpha_2) + \alpha_2 b$$

وباستخدام بعض العمليات الجبرية، يصبح تقدير المعلمتين $(a،b)$ كالتالي:

$$\hat{a} = \left[x_2\alpha_1 - x_1\alpha_2\right] / \left[\alpha_1 - \alpha_2\right]$$

$$\hat{b} = \left[x_2 - \hat{a}(1-\alpha_2)\right] / \alpha_2$$

المثال (٢-٤). على افتراض أن المحلل يريد استخدام التوزيع المنتظم المتصل على بعض البيانات حيث يعطي الخبراء قيم التقدير $x_1 = 10$ و $x_2 = 100$.

إذا كان $\alpha_1 = 0.0$ و $\alpha_2 = 1.0$، تصبح المعادلات المذكورة أعلاه معرفة كما يلي:

$$\hat{a} = \left[100 \times 0.0 - 10 \times 1.0\right] / \left[0.0 - 1.0\right] = 10$$

$$\hat{b} = \left[100 - 10\left(1-1\right)\right] / 1 = 100$$

وإذا كان $\alpha_1 = 0.0$ و $\alpha_2 = 0.8$، تصبح التقديرات:

$$\hat{a} = \left[100 \times 0.0 - 10 \times 0.8\right] / \left[0.0 - 0.8\right] = 10$$

$$\hat{b} = \left[100 - 10(1-0.8)\right] / 0.8 = 122.5$$

وإذا كان $\alpha_1 = 0.2$ و $\alpha_2 = 1.0$، تصبح التقديرات:

$$\hat{a} = \left[100 \times 0.2 - 10 \times 1.0\right] / \left[0.2 - 1.0\right] = -12.5$$

$$\hat{b} = \left[100 - (-12.5)(1.0-1.0)\right] / 1.0 = 100$$

وإذا كان $\alpha_1 = 0.2$ و $\alpha_2 = 0.8$، تصبح التقديرات:

$$\hat{a} = \left[100 \times 0.2 - 10 \times 0.8\right] / \left[0.2 - 0.8\right] = -20$$

$$\hat{b} = \left[100 - (-20)(1-0.8)\right] / 0.8 = 130$$

المثال (٢-٥). في نظام التصنيع تحتاج الإدارة إلى تقدير الزمن اللازم لتصميم منتج جديد. وعند الاستفسار من المهندسين سيكون تقدير منتصف المدى %50 من 100 إلى 120 ساعة. وهو ما يعني أنَّ النقطة %0.25 والنقطة %0.75 تكونان على النحو التالي:

$$x_{0.25} = 100$$

$$x_{0.75} = 120$$

وكذلك

$$\alpha_1 = 0.25$$

$$\alpha_2 = 0.75$$

لذلك يتم حساب تقدير قيم المعالم على النحو التالي:

$$\hat{a} = \left[120 \times 0.25 - 100 \times 0.75\right] / \left[0.25 - 0.75\right] = 90$$

$$\hat{b} = \left[120 - (90)(1 - 0.75)\right] / \ 0.75 = 130$$

ويكون تقدير الزمن اللازم لإنجاز هذا التصميم ما بين 90 إلى 130 ساعة.

٢-٦ ملخص

يتم استخدام التوزيع المنتظم المتصل عندما يقع المتغير العشوائي في أي مكان بين حدين، وغالباً ما يتم استخدام التوزيع عندما يكون لدى المحلل بعض المعلومات البسيطة عن التوزيع، ويلزم معرفة تقدير المعلمتين $(a، b)$ وعندما تتوفر بيانات العينة، يتم حساب قيم المعالم، إما من خلال طريقة الإمكان الأعظم وإما من خلال طريقة العزوم. وعند عدم توفر بيانات العينة، يتم استشارة أحد الخبراء للحصول على التقدير المناسب.

المثال (٢-٥). في نظام التصنيع تحتاج الإدارة إلى تقدير الزمن اللازم لتصميم منتج جديد. وعند الاستفسار من المهندسين سيكون تقدير منتصف المدى 50% من 100 إلى 120 ساعة. وهو ما يعني أنَّ النقطة 0.25% والنقطة 0.75% تكونان على النحو التالي:

$$x_{0.25} = 100$$

$$x_{0.75} = 120$$

وكذلك

$$\alpha_1 = 0.25$$

$$\alpha_2 = 0.75$$

لذلك يتم حساب تقدير قيم المعالم على النحو التالي:

$$\hat{a} = \left[120 \times 0.25 - 100 \times 0.75\right] / \left[0.25 - 0.75\right] = 90$$

$$\hat{b} = \left[120 - (90)(1 - 0.75)\right] / \; 0.75 = 130$$

ويكون تقدير الزمن اللازم لإنجاز هذا التصميم ما بين 90 إلى 130 ساعة.

٢-٦ ملخص

يتم استخدام التوزيع المنتظم المتصل عندما يقع المتغير العشوائي في أي مكان بين حدين، وغالباً ما يتم استخدام التوزيع عندما يكون لدى المحلل بعض المعلومات البسيطة عن التوزيع، ويلزم معرفة تقدير المعلمتين $(a، b)$ وعندما تتوفر بيانات العينة، يتم حساب قيم المعالم، إما من خلال طريقة الإمكان الأعظم وإما من خلال طريقة العزوم. وعند عدم توفر بيانات العينة، يتم استشارة أحد الخبراء للحصول على التقدير المناسب.

الفصل الثالث

التوزيع الأسّي
EXPONENTIAL DISTRIBUTION

يبلغ التوزيع الأسي ذروته عندما تكون قيمة المتغير العشوائي صفراً وينخفض تدريجياً مع ازدياد قيمة هذا المتغير. إنَّ للتوزيع الأسي معلمة واحدة وله حسابات سهلة بالنسبة لدالة الكثافة الاحتمالية ودالة التوزيع التراكمية. كما يتميّز التوزيع بخاصية فقدان الذاكرة حيث إن احتمالية حدوث الحدث التالي ضمن مدى ثابت هي ذاتها بغض النظر عن زمن بدء المدى. يعتبر هذا التوزيع من الأساسيات في نظرية صفوف الانتظار باعتبار أنه مستخدم كمتغير بالنسبة للزمن الممتد بين الوصول إلى النظام، كذلك زمن الخدمة. ولهذا التوزيع أيضاً تطبيقات في دراسة الموثوقية حيث يتم تحديدها كزمن الإخفاق بالنسبة لأحد العناصر، وعندما تكون قيمة المعلمة مجهولة، يتم استخدام بيانات العينة للوصول إلى تقدير لهذه المعلمة. وعند عدم توفر بيانات العينة، يُتيح القياس التقريبي للتوزيع للمحلل تقدير قيمة هذه المعلمة.

٣-١ أساسيات

إن للمتغير العشوائي الأسي X مدى مقبولاً من صفر وأكبر حيث تبلغ ذروة الحدوث في الصفر وتنخفض تدريجياً مع ازدياد x . وتعرف دالة الكثافة الاحتمالية $f(x)$ ذات المعلم الواحد θ كما يلي:

$$f(x) = \theta e^{-\theta x} \qquad x \geq 0$$

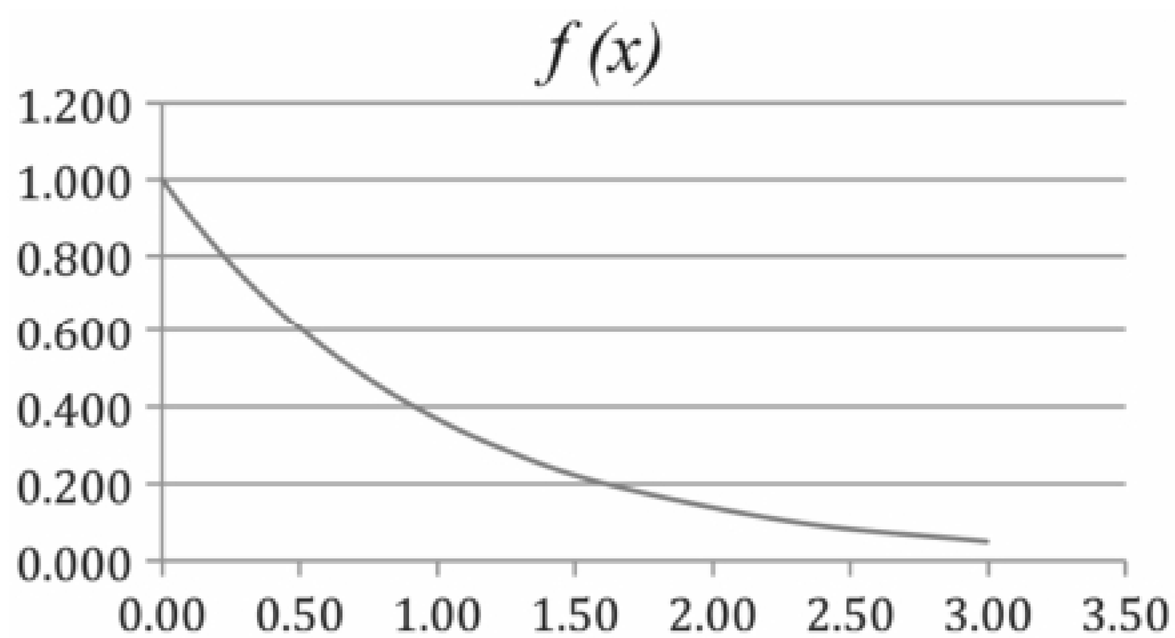
حيث تمثل هذه الدالة بيانياً كما في الشكل (٣-١).

ويمكن الحصول على المتوسط والتباين، والانحراف المعياري للمتغير X في الصور التالية:

$$\mu = \frac{1}{\theta}$$

$$\sigma^2 = \frac{1}{\theta^2}$$

$$\sigma = \frac{1}{\theta}$$



شكل (٣-١). الكثافة الاحتمالية للتوزيع الأسي عندما تكون θ =1.

ونظراً لأن المتوسط والانحراف المعياري متساويان، يكون معامل الاختلاف على النحو التالي:

معامل الاختلاف = $\frac{\sigma}{\mu} = 1$

كما تعرف دالة الاحتمال التراكمية للمتغير X على النحو التالي:

$$F(x) = 1 - e^{-\theta x} \qquad x \geq 0$$

نلاحظ أنه عندما : $x = x'$

$$F(x') = P(X \leq x')$$

ومن ذلك

$$F(0) = P(x' = 0) = 0$$

ويكون

$$F(\infty) = P(x` = \infty) = 1$$

تكون النقطة المئوية α في التوزيع، والمشار إليها بالرمز x_α، هي قيمة x حيث إن احتمال X الذي هو أقل أو يساوي x_α هو α، على النحو التالي:

$$P(X \le x_\alpha) = \alpha$$

يتم الحصول على قيمة x_α من خلال العلاقة:

$$x_\alpha = -\left(1 / \theta\right) Ln\left(1 - \alpha\right)$$

حيث Ln هو اللوغاريتم الطبيعي.

المثال (٣-١). يصل العملاء إلى متجر ما بمتوسط 5 دقائق وذلك من خلال التوزيع الأسي مع المتغير العشوائي X. وبالتالي

μ = 5 = متوسط الدقائق بين مرات الوصول

$\theta = 5\ /\ 1 = 0.2 =$ متوسط عدد مرات الوصول بالدقيقة

إن دالة الكثافة الاحتمالية للمتغير X هي على النحو التالي:

$$f(x) = 0.2e^{-0.2x} \qquad x \ge 0$$

ودالة التوزيع التراكمية هي:

$$F(x) = 1 - e^{-0.2x} \qquad x \ge 0$$

نلاحظ أنه على سبيل المثال، عندما $\alpha = 0.5$، نجد أن

$$x_{0.5} = \left(-1 / 0.2\right) 1\text{n} \left(1 - 0.5\right) = 3.47 \text{ دقيقة}$$

وبالتالي

$$F(3.47) = P(X \le 3.47) = 0.50$$

وكذلك عندما α=0.9 نجد أن

$$x_{0.9} = \left(-1 / 0.2\right) \ln\left(1 - 0.9\right) = 11.51$$ دقيقة

ومن ثم

$$F(11.51) = P(X \leq 11.51) = 0.90$$

٣-٢ القيم الجدولية

يبيّن الجدول (٣-١) مقارنة لقيم $x\alpha$ المساوية للنقطة المئوية α عندما تختلف α من 0.00 إلى 0.95، وتختلف θ من 0.1 إلى 10. كما يبيّن القيمة المقابلة لمتوسط X ، μ، بالنسبة لكل قيمة من قيم θ. نلاحظ أنه عندما تكون θ=1، فإن القيم الدنيا وحتى العليا في هذا المدى هي من 0 إلى 3. وعندما تكون $\theta > 1$، يزداد المدى، وعندما تكون $\theta < 1$، ينخفض المدى. يمثل الشكل (٣-١) التوزيع عندما تكون $\theta = 1$.

جدول (٣-١). يبيّن الجدول النقاط المئوية α،xα، بالنسبة للمعلمة الأسية $\theta = (0.1 - 10)$، $\alpha = (0 - 0.95)$ **و** $\mu = (0.1 - 10)$.

	θ	0.1	0.5	1	5	10
α	μ	10	2	1	0.2	0.1
0.00		0.00	0.00	0.00	0.00	0.00
0.05		0.51	0.10	0.05	0.01	0.01
0.10		1.05	0.21	0.11	0.02	0.01
0.20		2.23	0.45	0.22	0.04	0.02
0.30		3.57	0.71	0.36	0.07	0.04
0.40		5.11	1.02	0.51	0.10	0.05
0.50		6.93	1.39	0.69	0.14	0.07
0.60		9.16	1.83	0.92	0.18	0.09
0.70		12.04	2.41	1.20	0.24	0.12
0.80		16.09	3.22	1.61	0.32	0.16
0.90		23.03	4.61	2.30	0.46	0.23
0.95		29.96	5.99	3.00	0.60	0.30

٣-٣ خاصية فقدان الذاكرة

يتميّز التوزيع الأسي بخاصية فقدان الذاكرة وذلك لأن احتمال وقوع الحدث التالي بعد الفاصل الزمني x هو ذاته بغض النظر عن متى يحدث زمن بدء الفاصل الزمني. قد يكون حدوث زمن البدء مباشرةً بعد الحدث الأخير أو أي زمن بعد ذلك. لإظهار هذه العلاقة، نلاحظ فيما يلي احتمال أن يكون الزمن أكبر من x :

$$\begin{aligned} H(x) &= 1 - F(x) \\ &= 1 - \left[1 - e^{-\theta x}\right] \\ &= e^{-\theta x} \end{aligned}$$

على افتراض مرور الفاصل الزمني بطول x_0 بعد الحدث الأخير، وبافتراض احتمال وقوع الحدث التالي بعد طول الفاصل التالي x ، يكون الاحتمال الشرطي بأن زمن الوصول أكبر من x بالنسبة للحدث التالي وذلك بمعرفة نقطة البدء منذ الحدث الأخير $\left(x_0 + x\right)$ على النحو التالي:

$$\begin{aligned} H(x|x_0) &= H(x_0 + x) \;/\; H(x_0) \\ &= e^{-\theta\left(x_0 + x\right)} \;/\; e^{-\theta x_0} \\ &= e^{-\theta x} \end{aligned}$$

باعتبار أن

$$H(x) = H(x|x_0) = e^{-\theta x}$$

لذا فإن الاحتمال هو ذاته بغض النظر عن مكان نقطة البدء، وبالتالي فإن للتوزيع خاصية فقدان الذاكرة.

المثال (٣-٢). في المثال (٣-١)، متوسط الزمن بين مرات الوصول μ=5 دقائق، والمعلمة الأسية θ=0.2 مع ملاحظة أن احتمال x = 11.51 دقيقة أو أقل هو 0.90. لذلك فإن احتمال أن يصبح الحدث التالي أكبر من 11.51 هو:

$$P(X > 11.51) = 1 - F(11.51)$$
$$= 1 - 0.90$$
$$= H(11.51)$$
$$= 0.10$$

على افتراض أنه بعد انقضاء $x_0 = 2.0$ دقيقة يبدأ المحلل بحساب الزمن حتى وقوع الحدث التالي. ومن المفيد إيجاد احتمال وقوع الحدث التالي بعد 11.51 دقيقة. علماً بأن الزمن الذي يلي الحدث الأخير هو $2 + 11.51 = 13.51$ دقيقة ويتم حساب الاحتمال قيد البحث على النحو التالي:

$$H(13.51|2.00) = H(13.51) \;/\; H(2.00)$$
$$= e^{-\left(0.2\times 13.51\right)} \;/\; e^{-\left(0.2\times 2.00\right)}$$
$$= 0.067 \;/\; 0.670$$
$$= 0.10$$

وبالتالي

$$H(13.51|2.00) = H(11.51) = 0.10$$

٣-٤ علاقة بواسون

تم وصف توزيع بواسون في الفصل السابع عشر. إن المتغير العشوائي لتوزيع بواسون n هو عدد الأحداث التي ستقع خلال فاصل زمني معيّن عندما يكون معدّل الحدوث θ ثابتاً. وعندما يكون معلم التوزيع الأسي هو ذاته θ، فإن المتغير العشوائي X يمثل الزمن بين وقوع أحداث بواسون، لذا تلعب هذه العلاقة القائمة بين توزيعي بواسون والأسي دوراً مهماً في تطوير نظرية صفوف الانتظار.

احتمال n من توزيع بواسون يكون على النحو التالي:

$$P(n) = \frac{\theta^{n} e^{-\theta}}{n!} \qquad n = 0,1,2,\ldots$$

ودالة الكثافة الاحتمالية للتوزيع الأسي هي

$$f(x) = \theta e^{-\theta x} \qquad x \geq 0$$

كما أن القيمة المتوقعة لكل متغير من المتغيرين تكون على النحو التالي:

$$E(n) = \theta$$

$$E(x) = \frac{1}{\theta}$$

المثال (٣-٣). بالعودة للمثال (٣-١)، حيث إن X هو التوزيع الأسي ويمثل الدقائق بين مرات الوصول على أن يكون $\mu = 5.0$ والمعلمة $\theta = 0.20$. وباعتبار أن التوزيع الأسي مرتبط بتوزيع بواسون، يصبح المتغير العشوائي n والذي يمثل عدد مرات الوصول في الدقيقة الواحدة ويصبح احتمال n كما يلي:

$$P(n) = \frac{0.2^n e^{-0.2}}{n!} \qquad n = 0,1,2,\ldots$$

والقيمة المتوقعة لـ n هي:

$$E(x) = 0.2$$

نلاحظ أن

$$P(0) = \frac{0.2^0 e^{-0.2}}{0!} = 0.819$$

$$P(1) = \frac{0.2^1 e^{-0.2}}{1!} = 0.164$$

$$P(2) = \frac{0.2^2 e^{-0.2}}{2!} = 0.016$$

$$P(3) = \frac{0.2^3 e^{-0.2}}{3!} = 0.001$$

٣-٥ بيانات العينة

عندما يريد المحلل تطبيق التوزيع الأسي في دراسة ما، وتكون قيمة المعلمة θ مجهولة، يمكن استخدام بيانات العيّنة لتقدير قيمة هذه المعلمة. إن بيانات العينة هي n من المشاهدات والتي تم أخذها عشوائياً ويُشار إليها بواسطة $(x_1, \ldots, x_n)$. ولتطبيق التقدير تم أخذ الإحصائية التالية من بيانات العينة:

$\overline{x}$ = المتوسط

٣-٦ تقدير المعلمة من بيانات العينة

باستخدام متوسط العينة $\overline{x}$ يتم الحصول على تقدير قيمة المعلمة كما هو مبين أدناه:

$$\hat{\theta} = \frac{1}{\overline{x}}$$

المثال (٣-٤). في إحدى الدراسات التي تم إعدادها لتحديد تلوث الهواء في مطار محلي، جمع المحلل فترات بين وصول الطائرات $n=10$ كالتالي $0.5,1.7,2.2,1.9,4.3,8.3,0.4,1.8,3.8,2.7$. نورد فيما يلي المتوسط، والانحراف المعياري، ومعامل الاختلاف:

$$\overline{x} = 2.80$$
$$S = 2.27$$

$$COV = 0.81$$

نلاحظ أن معامل اختلاف العينة = 0.81 قريب بشكلٍ مقبول من معامل الاختلاف الأسي النظري والذي يساوي

$$\frac{\sigma}{\mu} = 1$$

ويصبح تقدير المعلم الأسي كما يلي:

$$\hat{\theta} = \frac{1}{\bar{x}} = \frac{1}{2.80} = 0.357$$

ودالة الكثافة الأسية المقدّرة للمتغير X على النحو التالي:

$$f(x) = 0.357e^{-0.357x} \qquad x \geq 0$$

٣-٧ تقدير المعلمة عند عدم توفر بيانات

عندما يريـد المحلل تطبيق التوزيع الأسي فـي إحـدى الدراسات، ولا يوجـد لديـه تقديـر للمعلمة θ وليس لديه أيضاً بيانات العينة، فعندئذ يطلب الـمحلل الاستشارة مـن أحـد الخبراء الذي يعطيه بـدوره تقريباً للنقطـة المئويـة α بالنسبـة للمتغير X والمُشار إليها بالرمز x_{α}. نلاحظ أن:

$$P(X \leq x_{\alpha}) = \alpha = 1 - e^{-\theta x_{\alpha}}$$

وبالتالي

$$1 - \alpha = e^{-\theta x_{\alpha}}$$

وباستخدام بعض العمليات الجبرية، يصبح تقدير قيمة المعلمة:

$$\hat{\theta} = -\left(1 / x_{\alpha}\right) Ln\left(1 - \alpha\right)$$

كما نلاحظ أيضاً أن متوسط X يعطى بالعلاقة:

$$\hat{\mu} = \frac{1}{\hat{\theta}}$$

المثال (٣-٥). في إحدى الدراسات التي تم إعدادها لتحديد المدة الزمنية التي يتم استغراقها لتنفيذ إحدى المهام حيث يتبع الاختلاف للتوزيع الأسي ولم يكن لدى المحلل البيانات اللازمة لتقدير قيمة المعلمة. وعند استشارته لأحد الخبراء فيما يتعلق بالنظام، يقدّر الخبير بأن 50% من المهام ستستغرق 10 دقائق أو أقل. باستخدام هذا التقريب، تم ملاحظة البيانات التالية:

$x_{0.50} = 10$ دقائق

$$\alpha = 0.50$$

ولذلك يصبح تقدير المعلم الأسي على النحو التالي:

$$\hat{\theta} = -\left(1 / 10\right) Ln \left(1 - 0.50\right) = 0.069$$

وبالتالي فإن دالة الكثافة الاحتمالية ودالة التوزيع التراكمية تكونان على النحو التالي:

$$f(x) = 0.069\, e^{-0.069x} \qquad x \geq 0$$

$$F(x) = 1 - e^{-0.069x} \qquad x \geq 0$$

المثال (٣-٦). تحدث أعطال في النظام الكهربائي من حين إلى آخر ويتم ارسال المهندسين لإجراء الإصلاحات اللازمة حيث يستغرق زمن الإصلاح حوالي 100 دقيقة أو أقل بنسبة تصل لـ 90% من الأعطال. على افتراض أن زمن إصلاح العطل يتبع التوزيع الأسي، لذا يتم التوصل إلى تقدير قيمة المعلمة كما هو مبيّن أدناه.

نلاحظ أولاً:

$$\alpha = 0.90$$

$x_{0.90} = 100$ دقيقة

إذاً

$$\hat{\theta} = -\left(1 / 100\right) Ln \left(1 - 0.90\right) = 0.023$$

$\hat{\mu} = \dfrac{1}{0.023} = 43.43$ دقيقة

وبالتالي

$$f(x) = 0.023 e^{-0.023x} \qquad x \geq 0$$

$$F(100) = 1 - e^{-0.023\left(100\right)} = 0.90$$

٣-٨ ملخص

إن للتوزيع الأسي معلمة واحدة وهي θ بحيث تبلغ دالة الكثافة الاحتمالية ذروتها عند الصفر وتنخفض تدريجياً بعد ذلك. إن للمتوسط والانحراف المعياري القيمة ذاتها، وبالتالي فإن معامل الاختلاف يكون مساوياً للواحد دائماً. وعندما تتوفر بيانات العينة، يتم حساب تقدير قيمة المعلمة بسهولة، وعند عدم توفر البيانات، يتم التوصل إلى تقريب للمعلمة باستخدام مقياس يوفره أحد الخبراء.

الفصل الرابع

توزيع إرلانج
ERLANG DISTRIBUTION

٤-١ مقدمة

ينسب توزيع إرلانج إلى أغنر إرلانج، وهو مهندس دنماركي يعمل في شركة هاتف كوبنهاغن، والذي كان يبحث عام 1908 عن عدد الوحدات اللازمة لاستيعاب الصوت في نظام الهاتف الخاص بهم. إن لتوزيع إرلانج معلمتين هما k وθ ، حيث إن k يمثل عدد المتغيرات الأسية التي تم اختصارها لتشكل متغير إرلانج. إن للمتغيرات الأسية المعلمة ذاتها θ كما هو الحال مع إرلانج. كما أن لتوزيع إرلانج أشكالاً تتراوح ما بين أسية وطبيعية. يُستخدم هذا التوزيع كثيراً في دراسة أنظمة صفوف الانتظار والتي تمثل الزمن المستغرق بين مرات الوصول، وكذلك الزمن المُستغرق لصيانة الوحدة. يبيّن هذا الفصل كيفية حساب دالة الاحتمال التراكمية $F(x)$، حيث إن X هو المتغير العشوائي. وعندما تكون قيم المعالم مجهولة وتتوفر بيانات العينة، تُستخدم قياسات من البيانات لتقدير قيم هذه المعالم. وعندما تكون قيم المعالم مجهولة ولا تتوفر بيانات العيّنة، يتم التوصل إلى تقريبات لبعض القياسات من التوزيع من قبل أحد الخبراء وهو ما يتيح لنا تقدير قيم المعالم المجهولة. يعتبر توزيع إرلانج حالة خاصة من توزيع جاما كما سيتضح لنا في الفصل الخامس، حيث إن المعلمة k هي عدد صحيح موجب بالنسبة لتوزيع إرلانج وهي أي عدد موجب بالنسبة لتوزيع جاما.

٤-٢ أساسيات

إن معلمات توزيع إرلانج هي على النحو التالي:

$\theta > 0$ = معلمة المقياس

معلمة الشكل = $k = (1, 2, \ldots)$

إن المعلمة k هي عدد صحيح موجب والمتغير العشوائي X هو مجموع المتغيرات العشوائية الأسية Y للمعلمة k مع ذات مقياس المعلمة θ. نلاحظ ما يلي:

$$X = Y_1 + \ldots + Y_k$$

حيث إن المتغير X يمثل توزيع إرلانج، والمتغير Y يمثل التوزيع الأسي. وعندما تكون $k = 1$، يكون توزيع إرلانج مطابقاً للتوزيع الأسي، وعندما تزداد قيمة k يقترب توزيع إرلانج من التوزيع الطبيعي نتيجة نظرية النهاية المركزية.

إن دالة الكثافة الاحتمالية ودالة التوزيع التراكمية تعرفان على النحو التالي:

$$f(x) = \frac{x^{k-1}\theta^k}{(k-1)!\, e^{-\theta x}} \qquad x > 0$$

$$F(x) = 1 - e^{-\theta x}\left[1 + (\theta x) + (\theta x)^2 / 2! + \ldots + (\theta x)^{k-1} / (k-1)!\right] \qquad x > 0$$

بينما يعرف المتوسط، والتباين، والانحراف المعياري على النحو التالي:

$$\mu = \frac{k}{\theta}$$

$$\sigma^2 = \frac{k}{\theta^2}$$

$$\sigma = \frac{\sqrt{k}}{\theta}$$

ومن ثم فإن معامل الاختلاف بالنسبة لتوزيع إرلانج يكون على النحو التالي:

معامل الاختلاف = $\frac{\sigma}{\mu} = \frac{1}{\sqrt{k}}$

نلاحظ أن قيم معامل الاختلاف متشابهه لقيم k وتتراوح ما بين 1 و9:

k	1	2	3	4	5	6	7	8	9
معامل الاختلاف	1.00	0.71	0.58	0.50	0.45	0.41	0.38	0.35	0.33

إن القيمة الأكثر تكراراً $\tilde{\mu}$ للمتغير X ، (المنوال) تكون على النحو التالي:

$$\tilde{\mu} = \frac{(k-1)}{\theta}$$

وعندما يكون k=1، فإن $\tilde{\mu} = 0$. كما نلاحظ أنه عندما تزداد قيمة k يزداد المنوال أيضاً.

٤-٣ الجداول

يبيّن الجدول (٤-١) قيم دالة التوزيع التراكمية $F(x)$ عندما تكون المعالم: $\theta = 1$ و $k=\left[1,\ (1),9\right]$. في حين أن مدى المتغير العشوائي هو: $\left[0,\ \left(0.5\right),18\right]$. كما يمكن استخدام جدول التوليد لإيجاد التوزيع التراكمي لأي توزيع من توزيعات إرلانج عندما لا تكون قيمة المعلمة $\theta = 1$. نورد فيما يلي كيفية تطبيق ذلك.

جدول (٤-١). توزيع إرلانج التراكمي، $F(x)$ ، عندما يكون θ=1 ، $x = 0$ - 18 ، و $k = 1$ - 9 .

x / k	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	0	0	0
0.5	0.39	0.09	0.01	0	0	0	0	0	0
1	0.63	0.26	0.08	0.02	0	0	0	0	0
1.5	0.78	0.44	0.19	0.07	0.02	0	0	0	0
2	0.86	0.59	0.32	0.14	0.05	0.02	0	0	0
2.5	0.92	0.71	0.46	0.24	0.11	0.04	0.01	0	0
3	0.95	0.80	0.58	0.35	0.18	0.08	0.03	0.01	0
3.5	0.97	0.86	0.68	0.46	0.27	0.14	0.07	0.03	0.01
4	0.98	0.91	0.76	0.57	0.37	0.21	0.11	0.05	0.02
4.5	0.99	0.94	0.83	0.66	0.47	0.30	0.17	0.09	0.04
5	0.99	0.96	0.88	0.73	0.56	0.38	0.24	0.13	0.07
5.5	1	0.97	0.91	0.80	0.64	0.47	0.31	0.19	0.11
6	1	0.98	0.94	0.85	0.71	0.55	0.39	0.26	0.15
6.5	1	0.99	0.96	0.89	0.78	0.63	0.47	0.33	0.21
7	1	0.99	0.97	0.92	0.83	0.70	0.55	0.40	0.27
7.5	1	1	0.98	0.94	0.87	0.76	0.62	0.48	0.34
8	1	1	0.99	0.96	0.90	0.81	0.69	0.55	0.41

تابع جدول (٤-١).

$x\ /\ k$	1	2	3	4	5	6	7	8	9
8.5	1	1	0.99	0.97	0.93	0.85	0.74	0.61	0.48
9	1	1	0.99	0.98	0.95	0.88	0.79	0.68	0.54
9.5	1	1	1	0.99	0.96	0.91	0.84	0.73	0.61
10	1	1	1	0.99	0.97	0.93	0.87	0.78	0.67
10.5	1	1	1	0.99	0.98	0.95	0.90	0.82	0.72
11	1	1	1	1	0.98	0.96	0.92	0.86	0.77
11.5	1	1	1	1	0.99	0.97	0.94	0.89	0.81
12	1	1	1	1	0.99	0.98	0.95	0.91	0.84
12.5	1	1	1	1	0.99	0.99	0.97	0.93	0.88
13	1	1	1	1	1	0.99	0.97	0.95	0.90
13.5	1	1	1	1	1	0.99	0.98	0.96	0.92
14	1	1	1	1	1	0.99	0.99	0.97	0.94
14.5	1	1	1	1	1	1	0.99	0.98	0.95
15	1	1	1	1	1	1	0.99	0.98	0.96
15.5	1	1	1	1	1	1	0.99	0.99	0.97
16	1	1	1	1	1	1	1	0.99	0.98
16.5	1	1	1	1	1	1	1	0.99	0.98
17	1	1	1	1	1	1	1	0.99	0.99
17.5	1	1	1	1	1	1	1	1	0.99
18	1	1	1	1	1	1	1	1	0.99

ليكن:

$$F\left(x\alpha|\theta,k\right) = P\left(X \leq x\alpha|\theta,k\right)$$

يبيّن الجدول (٤-١) قيم $F(x\alpha) = F\left(x\alpha|\theta = 1,k\right)$ عندما تكون θ=1 و k تساوي من 1 إلى 9. باستخدام القليل من الرياضيات، تحدث العلاقة التالية:

بملاحظة أنه عندما يكون θ=1، فإن $x\alpha$ هي النقطة المئوية α لـقيمة معيّنة k. ليكن ‘xα = النقطة المئوية α لأي θ و لتبقى قيمة k ثابتة. تكون العلاقة بين النقطتين المئويتين α على النحو التالي:

$$x\alpha' = \frac{x\alpha}{\theta}$$

نلاحظ أيضاً أن متغير إرلانج بالنسبة لأي مجموعة θ, x، متعلق بقيمة X الجدولية، على النحو التالي:

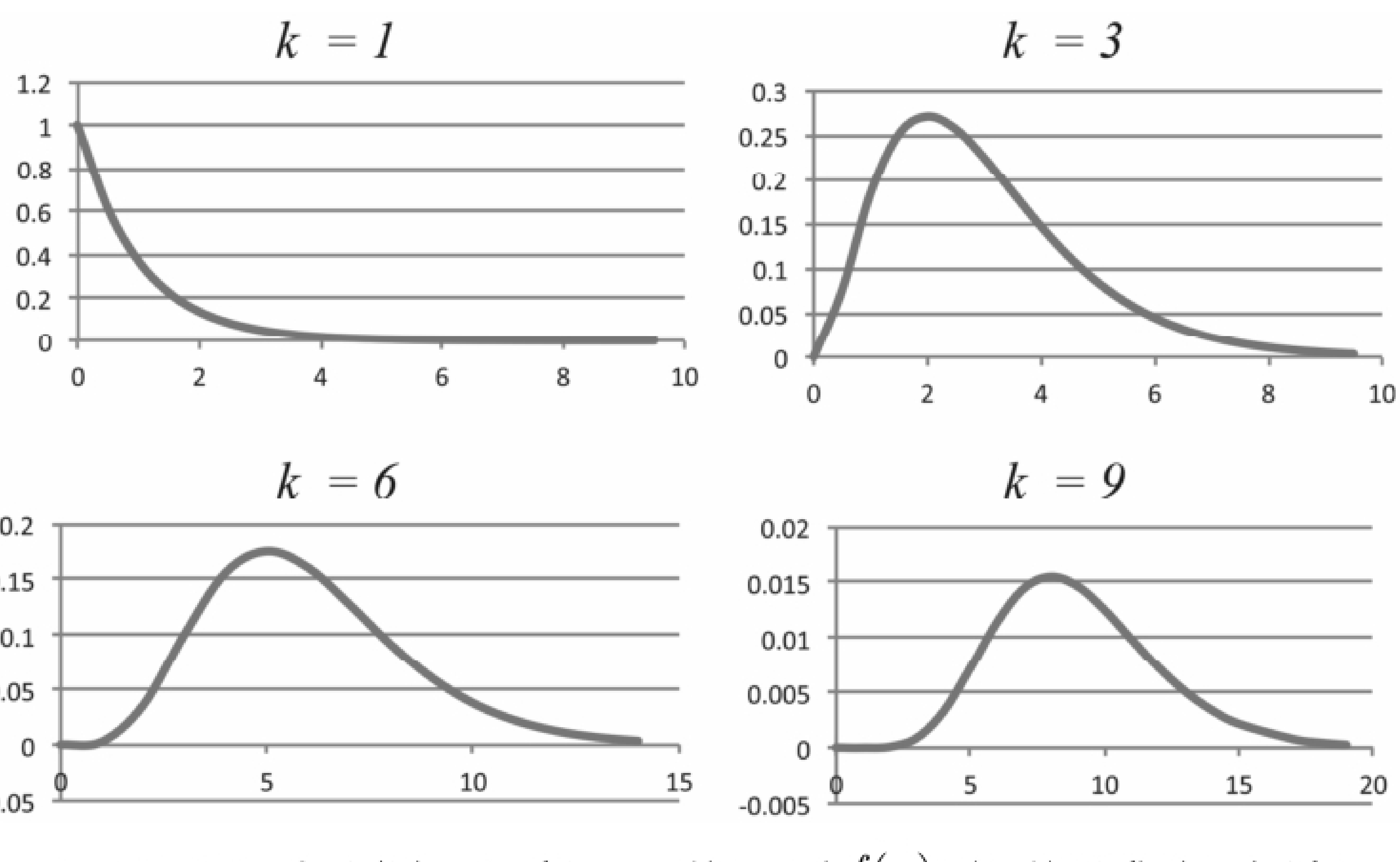
$$x' = \frac{x}{\theta}$$

المثال (٤-١). نلاحظ من الجدول (٤-١) حيث $F(5.0) = 0.73$ عندما يكون $\theta=1$ و $k=4$. إذا كان $\theta = 0.5$ تصبح النقطة 0.73% مساوية للمقدار $x'=5.0/0.5 = 10.0$، وبالتالي، $F(10.0) = 0.73$ عندما يكون $\theta = 0.5$ و $k= 4$.

هناك طريقة أخرى لحساب $F(x')$ عندما تكون $x'=10, \theta = 0.5$ و $k = 4$، وذلك من خلال ملاحظة $\theta x' = 0.5 \times 10 = 0.5$، وبالتالي:

$$F(10) = 1 - e^{-5}\left[1 + 5 + \frac{5^2}{2} + \frac{5^3}{6}\right] = 0.73$$

يمثل الشكل (٤-١) دالة الكثافة الاحتمالية $f(x)$، عندما تكون قيم المعالم $\theta=1$ و $k = 1, 3, 6, 9$. كما نلاحظ أنه عندما تكون المعلمة $k=1$، نحصل على التوزيع الأسي، وعندما تكون $k=9$، نحصل على التوزيع الطبيعي.



شكل (٤-١). الكثافة الاحتمالية $f(x)$ لتوزيع إرلانج، عندما تكون قيم المعالم $\theta=1$ و $k = 1, 3, 6, 9$.

٤-٤ بيانات العيّنة

عندما تتوفر بيانات $(x_1, x_2, \ldots, x_n)$ لعينة بحجم n، يتم قياس متوسط العينة، وتباينها وانحرافها المعياري. وتكون على النحو التالي:

$\overline{x}$ = متوسط العينة

S^2 = تباين العينة

S = الانحراف المعياري للعينة

٤-٥ تقدير المعلمتين عند توفر بيانات العينة

باستخدام قياسات العينة يمكن تقدير قيم المعلمتين من خلال طريقة العزوم. ولتحقيق ذلك، يتم استخدام العلاقة التي تربط بين المعالم ومتوسط وتباين المجتمع كما هو مبيّن سابقاً:

$$\mu = \frac{k}{\theta}$$

$$\sigma^2 = \frac{k}{\theta^2}$$

عند استبدال تقدير العينة المقابلة بالعلاقات المذكورة أعلاه، ينتج ما يلي:

$$\overline{x} = \frac{k}{\theta'}$$

$$S^2 = \frac{k'}{{\theta'}^2}$$

باستخدام بعض العمليات الجبرية، يكون تقدير المعالم على النحو التالي:

$$\theta' = \frac{\overline{x}}{S^2}$$

$$k' = \overline{x}\theta'$$

إلا أنه باعتبار أن k محصور في عدد صحيح موجب، يتم اختيار أقرب عدد صحيح إلى k ويتم تعديل تقدير المعالم وفقاً لذلك:

$\hat{k}$ = الحد الأدنى للمقدار $\left(k' + 0.5\right)$ والحد الأدنى للقيمة 1

$$\hat{\theta} = \frac{\hat{k}}{\overline{x}}$$

المثال (٤-٢). يوجد لدى مهندس يدرس زمن الإخفاق بالنسبة لإحدى المركبات التلقائية بيانات العينة التي تعطي المتوسط والتباين التاليين:

$$\overline{x} = 12 \text{ (100 - ساعة)}$$

$$S^2 = 20$$

على افتراض أنه قد تم توزيع المتغير العشوائي X بتوزيع إرلانج، يكون تقدير المعالم على النحو التالي:

$$\theta' = \frac{12}{20} = 0.60$$

$$k' = 12 \times 0.60 = 7.2$$

وعند تقريب k' إلى أقرب عدد صحيح له، ينتج تقدير معالم إرلانج بالصورة التالية:

$\hat{k}$ = الحد الأدنى $(0.5 + 7.2) = 7$

$$\hat{\theta} = \frac{7}{12} = 0.583$$

المثال (٤-٣). على افتراض أن المهندس المذكور في المثال السابق (٤-٢) يريد أيضاً تقدير النقطة 0.90% لزمن الإخفاق بالنسبة للمركبة. يتم تحقيق ذلك على النحو التالي: يتم بحث الجدول (٤-١) عند k=7 و $\theta = 1.0$ لإيجاد أقرب قيمة للمتغير X من خلال $F(x) = 0.90$. يُلاحظ ذلك في $x'_{0.90} = 10.5$. وتصبح القيمة المقابلة عندما تكون $\theta = 0.583$ و $k = 7$:

$$x_{0.90} = \frac{0.583}{10.5} = 18.5 \text{ (100 – ساعة)}$$

وبالتالي فإن

$$P(X \leq 18.0) = 0.90$$

٤-٦ تقدير المعالم عند عدم توفر بيانات

عند عدم توفر بيانات العينة مع تقديم أحد الخبراء تقدير عن شكل توزيع إرلانج، تكون قيم المعالم التي تم التوصل إليها بطريقة البحث، كما هو مبين أدناه.

$x\alpha$ = النقطة المئوية α

α = $F(x\alpha)$ = الاحتمال التراكمي

$\tilde{x}$ = القيمة المرجّحة (المنوال)

بالنسبة لكل قيمة من قيم k من 1 إلى 9، نوجد القيمة المقابلة للمعلمة θ من خلال العلاقة:

$$\tilde{x} = \frac{(k-1)}{\theta}$$

وهو ما يعطي:

$$\theta = \frac{(k-1)}{\tilde{x}}$$

بالنسبة لكل مجموعة من $(k،\theta)$ يتم تطبيق الدالة الاحتمالية التراكمية $F(x)$ المحددة مسبقاً لحساب $F(x\alpha)$. تعطي النتيجة 9 مجموعات من معلمات إرلانج التي تطابق بيانات الخبير، وتكون المجموعة التي تعطي $F(x\alpha)$ بالقرب من α هي تقدير المعالم المراد تطبيقها. في حال عدم قبول أي مجموعة فإن بيانات الخبير لا تشابه توزيع إرلانج.

المثال (٤-٤). يبيّن الجدول (٤-١) الحالات التي يتم فيها استخدام بيانات الخبير لتقدير قيم المعلم بالنسبة لتوزيع إرلانج. يتم توضيح معلومات الحالة في الجدول على

شكل: (1)، و(2)، و(3)، و(4). وتكون كل حالة عبارة عن بيانات الخبير الثلاثة التي تم تقديمها ($x\alpha$ ، α ، $\tilde{x}$). باستخدام بيانات الخبير، تتمثل الخطوة التالية في إيجاد وإدراج قيمة الاحتمال التراكمي $F(x\alpha)$ بالنسبة لكل قيمة من قيم k (من 1 إلى 9). حيث يتم حساب الاحتمال التراكمي $F(x)$ كما هو موضح سابقاً بالنسبة لأي x . كما تم إدراج القيمة المقابلة للمعلمة θ والتي يتم التوصل إليها من:

$$\theta = \frac{(k-1)}{\tilde{x}}$$

في الحالة الأولى تعطي مجموعة المعالم $(2,1) = (k,\theta)$:

$$F(3) = 1 - e^{-1\times 3}\left[1 + 1\times 3\right] = 0.801$$

وهي أقرب المجموعات التسع المرشحة إلى $\alpha = 0.80$، وبالتالي يتم اختيار المجموعة على أنها تقدير معالم إرلانج.

في الحالة الثانية تعطي المجموعة (6، 0.55):

$$F(17) = 1 - e^{-0.11\times 17}\left[1 + 0.11\times 17 + \ldots + 0.11\times 17^5\ /\ 5!\right] = 0.909$$

بالنسبة لأقرب المجموعات إلى $\alpha = 0.90$، وبالتالي يصبح تقدير المعالم لتوزيع إرلانج.

في الحالة الثالثة تعطي المجموعة (2،5)

$$(k-1)\theta = (5-1)(2) = 8$$

وبالتالي:

$$F(4) = 1 - e^{-8}\left[1 + 8 + 8^2\ /\ 2 + 8^3\ /\ 6 + 8^4\ /\ 24\right] = 0.900$$

بالنسبة لتوزيع إرلانج.

في الحالة الرابعة لا توجد مجموعة قريبة من $\alpha = 0.90$، وهو ما يشير إلى أن بيانات الخبير ليست مناسبة لتوزيع إرلانج (الجدول ٤-٢).

جدول (٤-٢). الحالات الأربع: (1-4) لبيانات الخبير التابعة لتوزيع إرلانج: ($x\alpha$ ، α ، $\tilde{x}$) التي تعطي $F(x\alpha)$ عندما تكون k من 1 إلى 9؛ والقيمة المقابلة $\theta = \dfrac{(k-1)}{\tilde{x}}$.

(1)

$x\alpha$	α	$\tilde{x}$							
3	0.8	1							
k	1	2	3	4	5	6	7	8	9
θ	0	1	2	3	4	5	6	7	8
$F(x\alpha)$	0	0.801	0.938	0.979	0.992	0.997	0.999	1.000	1.000

(2)

$x\alpha$	α	$\tilde{x}$							
17	0.9	9							
k	1	2	3	4	5	6	7	8	9
θ	0	0.11	0.22	0.33	0.44	0.55	0.66	0.77	0.88
$F(x\alpha)$	0	0.563	0.727	0.816	0.872	0.909	0.934	0.952	0.965

(3)

$x\alpha$	α	$\tilde{x}$							
4	0.9	2							
k	1	2	3	4	5	6	7	8	9
θ	0	0.5	1	1.5	2	2.5	3	3.5	4
$F(x\alpha)$	0	0.594	0.762	0.849	0.900	0.933	0.954	0.968	0.978

(4)

$x\alpha$	α	$\tilde{x}$							
4	0.9	3							
k	1	2	3	4	5	6	7	8	9
θ	0	0.33	0.67	1.00	1.33	1.67	2.00	2.33	2.67
$F(x\alpha)$	0.000	0.385	0.498	0.567	0.616	0.655	0.687	0.714	0.737

٤-٧ ملخص

إن لتوزيع إرلانج ذي المعلمتين k،θ عدة أشكال تتراوح ما بين أسية إلى طبيعية. ويُستخدم هذا التوزيع بشكلٍ كبير في دراسات صفوف الانتظار. وهو مشابه لتوزيع جاما عندما تكون المعلمة k عدداً صحيحاً موجباً. يتم تشكيل متغير إرلانج من خلال k من مجموع المتغيرات الأسية التي تشترك بذات المعلمة θ. وعندما تكون المعلمة $k = 1$ ، يكون التوزيع مشابهاً للتوزيع الأسي، وعند ازدياد k ونتيجة لنظرية النهاية المركزية، تقترب التوزيعات من الشكل الطبيعي. عندما يتم التوصل إلى بيانات العينة، يتم استخدام البيانات لتقدير قيم معالم إرلانج. وعندما لا تتوفر بيانات العينة، يتم استخدام القياسات التقريبية في التوزيع لتقدير قيمة هذه المعالم.

الفصل الخامس

توزيع جاما
GAMA DISTRIBUTION

٥-١ مقدمة

قدم كارل بيرسون (Karl Pearson)، وهو أستاذ بريطاني معروف، توزيع جاما في عام 1895. وتمت إعادة تسمية التوزيع الذي كان يسمى في الأصل التوزيع الثالث من نوع بيرسون في ثلاثينيات القرن الماضي بتوزيع جاما. هناك عدة أشكال لتوزيع جاما، وهي تتراوح ما بين شبه أسي إلى شبه طبيعي. كما أن للتوزيع معلمين وهما k وθ، حيث إن كليهما أكبر من الصفر. عندما يكون k عدداً صحيحاً موجباً، يكون التوزيع مشابهاً لتوزيع إرلانج. كما أنه عندما تكون k أقل من واحد أو مساوٍ له، يكون المنوال صفراً ويكون التوزيع شبه أسي. وعندما تكون k أكبر من واحد، يكون المنوال أكبر من صفر. وعند ازدياد k يكون الشكل مشابهاً للتوزيع الطبيعي. ولا يوجد حل مقرب لحساب الاحتمال التراكمي، إلا أنه قد تم تطوير طرق كمّية لإيجاده. كما تم تطوير طريقة أخرى في هذا الفصل وتُطبق عندما يتراوح مدى k ما بين 1 و9. عندما تتوفر بيانات العينة، يتم التوصل إلى تقدير قيم المعالم. وعند عدم توفر بيانات العينة، يتم التوصل إلى تقدير قيم المعالم باستخدام التقريبات في بعض قياسات التوزيع.

٥-٢ أساسيات

تكون معالم توزيع جاما على النحو التالي:

$\theta > 0$ = معلمة المقياس

$k > 0$ = معلمة الشكل

وتعرف دالة الكثافة الاحتمالية $f(x)$ على النحو التالي:

$$f(x) = x^{k-1}\theta^k e^{-\theta x} \big/ \Gamma(k), \qquad x > 0$$

٥-٣ دالة جاما

إن $\Gamma(k)$ هي دالة جاما، وهي ليست توزيع ويتم حسابها كما هو مبين أدناه:

$$\Gamma(0.5) = \sqrt{\pi}$$

$\Gamma(k) = (k-1)!$ حيث إن k عدد صحيح موجب

$$\Gamma(k) = \int_0^\infty w^{k-1} e^{-w} dw, \quad k > 0$$

بالنسبة للقيم الأخرى لـ $k > 1$، يمكن تطبيق صيغة ستيرلينغ [Abramowitz and Stegun (1964)، الصفحة 257] المدرجة أدناه:

$$\Gamma(k) = k^{k-0.5} e^{-k} \sqrt{2\pi} \left[1 + 1 / 12k + 1 / 288k^2 - 139 / 51840k^3 - 571 / 2488320k^4 + \cdots\right]$$

٥-٤ الاحتمال التراكمي

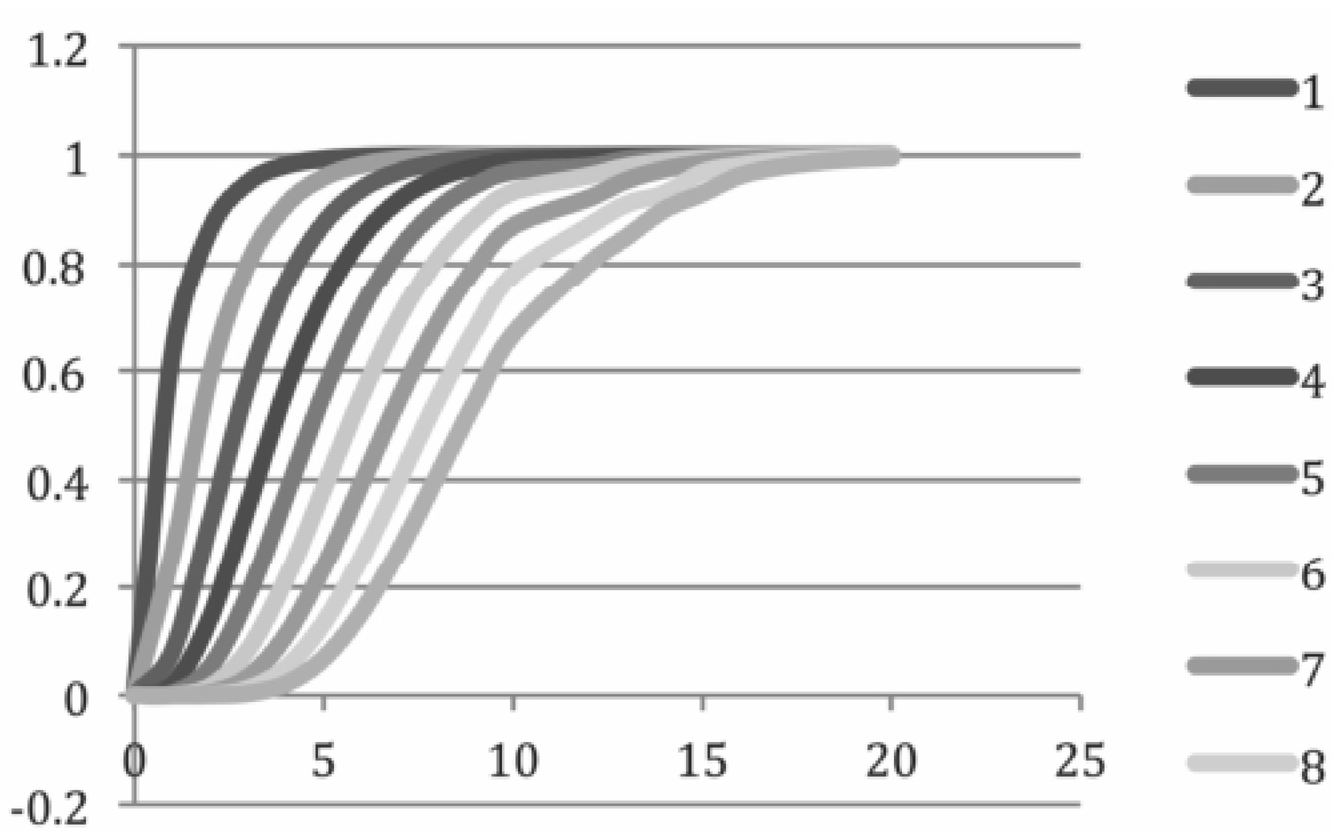
لا يوجد حل تقريبي لإيجاد دالة توزيع الاحتمال التراكمي لتوزيع جاما. إلا أنه عندما يكون k عدداً صحيحاً موجباً، يكون التوزيع مشابهاً لتوزيع إرلانج كما هو مبيّن في الفصل الرابع. في هذه الحالة الخاصة، يتم حساب دالة الاحتمال التراكمي كما هو مبيّن أدناه:

$$F(x) = 1 - e^{-\theta x} \left[1 + (\theta x) + (\theta x)^2 \big/ 2! + \cdots + (\theta x)^{k-1} \big/ (k-1)!\right], x > 0$$

يبيّن الجدول (٤-١) قيم $F(x)$ بالنسبة لمدى x المختار مع الأعداد الصحيحة للمعلمة k التي تتراوح ما بين 1 و9. يمثل الشكل (٥-١) شكل التوزيعات حيث إن $k = 1$ هو المنحنى على الجانب الأيسر، وتزداد k بشكلٍ مستمر إلى أن تصبح $k = 9$ المنحنى على الجانب الأيمن للشكل.

ويكون المتوسط، والتباين، والانحراف المعياري على النحو التالي:

$$\mu = k\big/\theta$$
$$\sigma^2 = k\big/\theta^2$$
$$\sigma = \sqrt{k}\big/\theta$$



شكل (٥-١). توزيع الاحتمال التراكمي بالنسبة لتوزيع جاما عندما يكون *k* عدداً صحيحاً يتراوح بين 1 و 9 من اليسار إلى اليمين.

كما يعرف معامل الاختلاف على النحو التالي:

معامل الاختلاف $= \sigma\big/\mu = 1\big/\sqrt{k}$

وتكون قيمة x الأكثر تكراراً (المنوال) على النحو التالي:

$\tilde{\mu} = (k-1)\theta$ عندما يكون $k \geq 1$

$\tilde{\mu} = 0$ عندما يكون $k < 1$

عندما يكون $k \leq 1$ يكون المنوال صفراً وشكل التوزيع شبه أسي، وعندما يكون $k > 1$ يكون المنوال أكبر من الصفر، وعندما يزداد k يكون التوزيع شبه طبيعي.

المثال (٥-١). على افتراض أنه كان لدى المحلل البيانات التي يكون فيها المتغير X يتبع توزيع جاما بحيث يكون $k = 3.0$ و $\theta = 0.8$. لذا تصبح دالة الكثافة الاحتمالية:

$$f(x) = 0.256x^2e^{-0.8x}, \qquad x > 0$$

لقد تم إدراج القيم المختارة للمتغير X (من 0 إلى 12) في الجدول (٥-١). وباعتبار أن k هو عدد صحيح، يمكن حساب الاحتمال التراكمي باستخدام $F(x)$ كما هو مبيّن في فصل توزيع إرلانج.

عندما يكون $k = 3$ و $x = 5$ ، فإن،

$$F(5) = 1 - e^{-4}\left[1 + 4 + 4^2/2\right] = 0.76$$

ويكون متوسط x ، وتباينه، وانحرافه المعياري على النحو التالي:

$$\mu = 3/0.8 = 3.75$$
$$\sigma^2 = 3/0.64 = 4.69$$
$$\sigma = 2.16$$

جدول (٥-١). دالة الكثافة الاحتمالية $f(x)$ بالنسبة للقيم المختارة للمتغير X الذي يتبع توزيع جاما حيث $k = 3$ و $\theta = 0.8$.

x	$f(x)$	x	$f(x)$
0	0	7	0.046
1	0.115	8	0.027
2	0.207	9	0.015
3	0.209	10	0.009
4	0.167	11	0.005
5	0.117	12	0.002
6	0.076		

ومن ثم فإن القيمة المرجّحة (المنوال) هي:

$$\tilde{\mu} = (3 - 1) \times 0.8 = 1.6$$

٥-٥ تقدير الاحتمال التراكمي

على افتراض أن قيمتي معالم جاما هما (k, θ) والهدف هو تقدير الاحتمال التراكمي لقيمة x' بمعلومية k و θ والمشار إليها هنا بـ $F\left(x' \middle| k, \theta\right)$.

باعتبار أن $x‘ = x/\theta , x = x‘\theta$ هي قيمة المتغير المقابل المرتبطة بمدخلات الجدول (٤-١). وبافتراض أن:

$$k_1 < k < k_2$$
$$x_1 < x < x_2$$

حيث إن k_1, k_2, x_1, x_2 هي أقرب مدخلات لـ k و x في الجدول (٤-١).

ولتسهيل الأمر نفرض أن $F\left(x \middle| k\right)$ ممثلاً للاحتمال التراكمي للمتغير X عندما تكون المعلمتان k و $\theta = 1$. إن قيم الاحتمال التراكمي $F\left(x \middle| k\right)$ المبينة في الجدول (٤-١) تكون على النحو التالي:

$$F\left(x_1 \middle| k_1\right)$$
$$F\left(x_2 \middle| k_1\right)$$
$$F\left(x_1 \middle| k_2\right)$$
$$F\left(x_2 \middle| k_2\right)$$

ومن ثم يتم التوصل إلى التقديرات الاحتمالية التالية:

$$F\left(x_1 \middle| k\right) \approx F\left(x_1 \middle| k_1\right) + (k - k_1)/(k_2 - k_1)\left[F\left(x_1 \middle| k_2\right) - F\left(x_1 \middle| k_1\right)\right]$$
$$F\left(x_2 \middle| k\right) \approx F\left(x_2 \middle| k_1\right) + (k - k_1)/(k_2 - k_1)\left[F\left(x_2 \middle| k_2\right) - F\left(x_2 \middle| k_1\right)\right]$$

وتكون

$$F\left(x \middle| k\right) \approx F\left(x_1 \middle| k\right) + (x - x_1)/(x_2 - x_1)\left[F\left(x_2 \middle| k\right) - F\left(x_1 \middle| k\right)\right]$$

وأخيراً، بالنسبة لـ $x' = x/\theta$ ، يصبح الاحتمال التراكمي:

$$F(x'/k, \theta) \approx F(x/k)$$

المثال (٥-٢). على افتراض أنه لدى المحلل بيانات جاما وقيم المعالم $k = 5.7$ و $\theta = 0.8$. يسعى المحلل لتقدير دالة الاحتمال التراكمي للمتغير $x‘$ الذي هو أقل أو يساوي 12. مع ملاحظة أن

$x = x'\theta = 12 \times 0.8 = 9.6$. ولتحقيق ذلك، يتم البحث في بيانات الجدول (٤-١) لإيجاد المدخلات الأقرب إلى $k = 5.7$ و $x = 9.6$ كما هو مبيّن أدناه:

$$\begin{aligned} k_1, k_2 &= 5, 6 \\ x_1, x_2 &= 9.5, 10 \end{aligned}$$

بعد ذلك تكون قيم الاحتمال ذات الصلة على النحو التالي:

$$\begin{aligned} F(9.5, 5) &= 0.96 \\ F(10.0, 5) &= 0.97 \\ F(9.5, 6) &= 0.91 \\ F(10.0, 6) &= 0.93 \end{aligned}$$

ومن ثم ينتج ما يلي

$$\begin{aligned} F\left(9.5 \mid 5.7\right) &\approx 0.96 + (5.7 - 5) \big/ (6 - 5) \left[0.91 - 0.96\right] = 0.925 \\ F(10.0 \mid 5.7) &\approx 0.97 + (5.7 - 5) \big/ (6 - 5) \left[0.93 - 0.97\right] = 0.942 \end{aligned}$$

وتكون

$$F(9.6 \mid 5.7) \approx 0.925 + (9.6 - 9.5) \big/ (10.0 - 9.5) \left[0.942 - 0.925\right] = 0.9284$$

وبالتالي فإن

$$F(x' = 12 \big| 9.6, 5.7) = P(x' < 12) = 0.925$$

٥-٦ بيانات العيّنة

على افتراض أن بيانات العينة المتاحة هي $(x_1, \ldots, x_n)$، تكون الاحصائيات التي تم قياسها من البيانات هي المتوسط، والتباين، والانحراف المعياري على النحو التالي:

$\overline{x} =$ متوسط العينة

$s^2 =$ تباين العينة

$s =$ الانحراف المعياري للعينة

٥-٧ تقدير المعلمتين عند توفر بيانات العينة

باستخدام مقاييس العينة يمكن تقدير قيم المعلمتين من خلال طريقة العزوم. ولتحقيق ذلك، يمكن استخدام العلاقة القائمة بين المعالم ومتوسط وتباين المجتمع كما هو محدد مسبقاً:

$$\mu = k/\theta$$
$$\sigma^2 = k/\theta^2$$

عند استبدال تقدير العينة المقابلة بالعلاقات المذكورة أعلاه ينتج ما يلي:

$$\overline{x} = \hat{k}/\hat{\theta}$$
$$s^2 = \hat{k}/\hat{\theta}^2$$

وباستخدام العمليات الجبرية، يكون تقدير المعالم كما هو مبين أدناه:

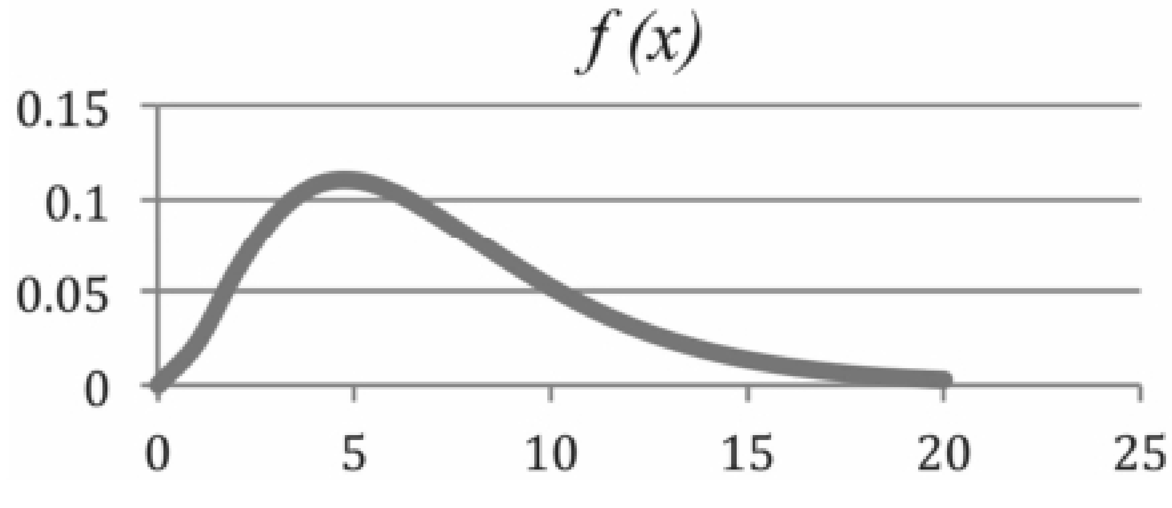
$$\hat{\theta} = \overline{x}/s^2$$
$$\hat{k} = \overline{x}\hat{\theta}$$

المثال ٥-٣: على افتراض أنه لدى المحلل بيانات العينة التالية ويريد تطبيق توزيع جاما:

(5.1,7.2,11.8,3.1,7.4,15.4,2.1,6.4,7.3,4.5)

وبالتالي فإن متوسط العينة، وتباينها، وانحرافها المعياري يكون على النحو التالي:

$$\overline{x} = 7.03$$
$$s^2 = 15.925$$
$$s = 2.236$$



شكل (٥-٢). توزيع جاما عندما يكون $k = 3.10$ و $\theta = 0.44$.

المثال (٥-٤). باستخدام بيانات المثال (٥-٣)، يتم حساب تقدير معلمتي جاما على النحو التالي:

$$\hat{\theta} = \overline{x}\big/s^2 = 7.03\big/15.925 = 0.44$$
$$\hat{k} = \overline{x}\hat{\theta} = 7.03 \times 0.44 = 3.10$$

حيث تم تمثيل الرسم البياني لتوزيع جاما في الشكل (٥-٢).

٥-٨ تقدير معلمة عند عدم توفر بيانات

عند عدم توفر بيانات العينة، يتم تقديم تقريب لشكل توزيع جاما. وقد تم توضيح قيم المعلمتين التي تم التوصل إليها عبر طريقة البحث أدناه. إن بيانات التقريب اللازمة هي على النحو التالي:

$x\alpha$ = النقطة المئوية α

$F(x\alpha) = \alpha =$ الاحتمال التراكمي

$\tilde{x} =$ القيمة المرجّحة (المنوال)

تفترض الطريقة المبيّنة هنا وقوع قيمة k في مكان ما بين (1 إلى 9). ولذلك بالنسبة لكل قيمة من قيم k التي تتراوح ما بين 1 و9، نوجد القيمة المقابلة للمعلمة θ من خلال العلاقة:

$$\theta = (k-1)\big/\tilde{x}$$

لكل مجموعه من (k,θ)، يتم تطبيق الدالة الاحتمالية $F(x)$ المحددة من خلال توزيع إرلانج لحساب $F(x\alpha)$، حيث تعطي النتيجة تسعة احتمالات تراكمية بالنسبة للمجموعتين (k,θ). يتم تحديد المجموعتين اللتين تعطيان أقرب احتمالات إلى α، حيث إن:

$$F(x\alpha\big|k_1,\theta) \le \alpha \le F(x\alpha\big|k_2,\theta)$$

و $k_2 = k_1 + 1$.

لذا يكون تقدير k الذي تم التوصل إليه على النحو التالي:

$$\hat{k} = k_1 + \Big[\alpha - F(x\alpha\big|k_1,\theta)\big/F(x\alpha\big|k_2,\theta) - F(x\alpha\big|k_1,\theta)\Big](k_2 - k_1)$$

ومن ثم يتم حساب القيمة المقابلة θ، والمشار إليها بـ $\hat{\theta}$، كما هو مبين أدناه:

$$\hat{\theta} = (\hat{k} - 1)\big/\tilde{x}$$

وأخيراً فإن تقدير المعالم هي:

$$(\hat{k}, \hat{\theta})$$

إذا لم تُقبل أي مجموعة فإن بيانات التقريب لا تماثل توزيع جاما.

المثال (٥-٥). يتم تصميم محاكاة مع الحاجة إلى متغير جاما عند عدم توفر بيانات العينة لتقدير قيم المعالم. ويتم استدعاء أحد الخبراء ليقوم بتقريب القياسات التالية في التوزيع: ستنخفض ما نسبته 90% من المشاهدات إلى ما دون 100، وتكون القيمة المرجحة 40. وبالتالي تكون القياسات على النحو التالي:

$$x_{0.9} = 100$$
$$\alpha = 0.90$$
$$\tilde{x} = 40$$

لتقدير قيم المعالم، يحدد المحلل أن $k = 1 - 9$، وبالنسبة لكل قيمة يتم قياس المعلمة θ، وكذلك دالة الاحتمال التراكمي لما هو أقل من 100 أو مساوياً لها $F(100)$. نبين فيما يلي كيفية حساب $F(100)$ عندما يكون $x = 100, k = 3$، و $\theta = 0.05$. ويلاحظ أن $\theta x = 0.05 \times 100 = 5.0$. وبالتالي:

$$F(100) = 1 - e^{-5}\left[1 + 5 + 5^2\big/2\right] = 0.875$$

حيث تم إدراج كافة النتائج أدناه:

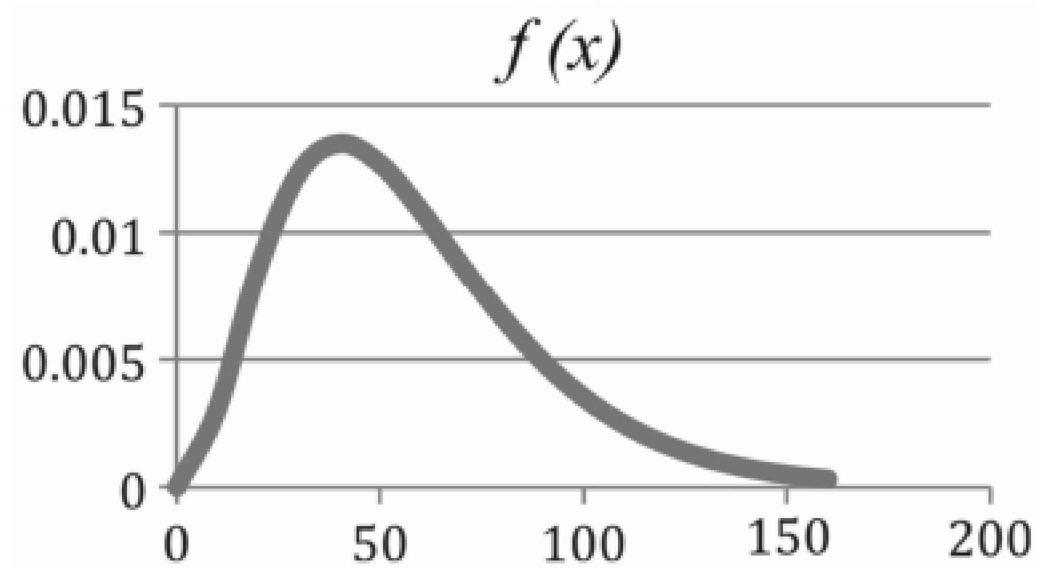
k	1	2	3	4	5	6	7	8	9
θ	0	.025	.050	.075	.100	.125	.150	.175	.200
$F(100)$	0	.717	.875	.941	.971	.985	.992	.996	.998

باعتبار أن $\alpha = 0.90$ يقع بين $k = 3, 4$ و $k_1 = 3$ و $k_2 = 4$، يتم التوصل إلى تقدير k كما هو مبين أدناه:

$$\hat{k} \sim 3 + \left[0.900 - 0.875\right] \Big/ \left[0.941 - 0.875\right] = 3.38$$

وتصبح القيمة المقابلة لـ θ :

$$\hat{\theta} = (\hat{k} - 1) \Big/ \tilde{x} = 2.38 \big/ 40 = 0.059$$



شكل (٥-٣). دالة كثافة جاما عندما تكون قيم المعالم $k = 3.38$ و $\theta = 0.059$.

وبالتالي تكون دالة كثافة احتمال جاما على النحو التالي:

$$f(x) = x^{2.38} (0.059)^{3.38}\, e^{(-0.059x)} \Big/ \Gamma(3.38)$$

وقد تم تمثيل الرسم البياني لدالة جاما في الشكل (٥-٣).

٥-٩ ملخص

يتّخذ توزيع جاما عدة أشكال، وتتراوح ما بين شبه أسي إلى طبيعي. لا يوجد في التوزيع صيغة حل مقربة بالنسبة للاحتمال التراكمي، إلا أن الطرق الكميّة متاحة للاستخدام. وقدم الفصل طريقة كمّية أخرى فيما يتعلق بحساب الاحتمال التراكمي. وعند توفرها، يتم جمع بيانات العينة لتقدير قيم المعالم. وعند عدم توفر بيانات العينة، يقدم الخبير قياسات التقريب التي تتيح تقدير قيم هذه المعالم.

الفصل السادس

توزيع بيتا
BETA DISTRIBUTION

٦-١ مقدمة

قدم عالم الرياضيات البريطاني كارل بيرسون توزيع بيتا في عام 1895، وكان يُسمى في الأصل بتوزيع بيرسون من النوع الأول. تم تغيير الاسم في أربعينيات القرن الماضي ليصبح توزيع بيتا. كما طبّق توماس بايز التوزيع عام 1763 كتوزيع لاحق لمعلمة توزيع برنولي. يتّخذ توزيع بيتا عدة أشكال تتراوح ما بين أسي، وأسي عكسي، ومثلثي أيمن، ومثلثي أيسر، وانحراف نحو اليمين، وانحراف نحو اليسار، وطبيعي، و شكل حوض الاستحمام. و العيب الوحيد في هذا التوزيع هو أنه من الصعب قليلاً استخدامه على تطبيقات حقيقية. إنَّ لتوزيع بيتا معلمتين أساسيتين هما k_1 وk_2 اللتان هما أكبر من الصفر؛ ومعلمتي موقع وهما a وb اللتان تعرّفان حدود المدى المقبول. وعندما تكون كلا المعلمتين أكبر من واحد، ينحرف متغير بيتا نحو اليسار أو نحو اليمين. وهي أشكال التوزيع الأكثر استخداماً ولهذا السبب يركّز الفصل على هذه الأشكال. ويُشار إلى المتغير العشوائي على أنه W حيث إن $(a \leq W \leq b)$. إن للمتغير X المرتبط بالمتغيرW، والذي يُسمى ببيتا المعياري مدى يتراوح ما بين 0 و1، حيث يتم تحديد الخصائص الرياضية لدالة الكثافة الاحتمالية من خلال توزيع بيتا المعياري. ويتضمن ذلك دالة بيتا ودالة جاما اللتين يصعب حسابهما. ولا توجد صيغة رياضية محددة بالنسبة لدالة الاحتمال التراكمي وبالتالي يتم توضيح الطريقة الكمية في أمثلة الفصل. كما يتم استخدام عملية للتحويل من W إلى X ومن X إلى W. وعند توفر بيانات العينة، يكون هناك حاجة إلى متوسط ومنوال العينة لتقدير قيمتي المعلمتين k_1 وk_2. ومن ثم وضع الانحدار لتقدير

متوسط X من منوال X. وأخيراً عند عدم توفر بيانات العينة، يتم التوصل إلى أفضل تقدير للحدين (a، b) والقيمة المرجّحة للمتغير W لتقدير قيم هذه المعالم.

٦-٢ أساسيات

يتم إدراج المتغير العشوائي لتوزيع بيتا على أنه W ويتضمن المدى (a إلى b)، في حين أن المدى للمتغير العشوائي المعياري المقابل X يتراوح ما بين (0 إلى 1). وتكون معالم بيتا على النحو التالي:

$(k_1, k_2) =$ المعلمتان

$(a, b) =$ المدى لمتغير بيتا W

$(0,1) =$ المدى لمتغير بيتا المعياري X

إن العلاقة بين W و X هي على النحو التالي:

$$x = (w-a)\big/b-a$$
$$w = a + x(b-a)$$

٦-٣ توزيع بيتا المعياري

إن دالة الكثافة الاحتمالية لتوزيع بيتا المعياري تعرف على النحو التالي:

$$f(x) = x^{k_1-1}\,(1-x)^{k_2-1}\Big/B(k_1,k_2)$$

حيث إن المقام هو دالة بيتا وليس التوزيع، ويتم حسابه على النحو التالي:

$$B(k_1,k_2) = \Gamma(k_1)\Gamma(k_2)\Big/\Gamma(k_1+k_2) \qquad 0 < x < 1$$

كما إن مركبات دالة بيتا هي دالة جاما كما هو مبيّن في الفصل الخامس.

ويكون متوسط وتباين توزيع بيتا على النحو التالي:

$$\mu = k_1\big/(k_1+k_2)$$
$$\sigma^2 = k_1k_2\Big/\left[(k_1+k_2)^2(k_1+k_2+1)\right]$$

وعندما يكون $k_1 > 1$ و $k_2 > 1$، تكون القيمة المرجّحة (المنوال) على النحو التالي:

$$\tilde{\mu} = (k_1 - 1)\big/(k_1 + k_2 - 2)$$

ونشير هنا بأنه لا يوجد حل محدد بالنسبة لدالة الاحتمال التراكمي $F(x)$، إلا أنه قد تم توضيح الطريقة الكمّية التي تعطي التقدير في الأمثلة القادمة.

٦-٤ أشكال توزيع بيتا

يتّخذ توزيع بيتا عدة أشكال اعتماداً على قيمة المعلمتين (k_1, k_2) حيث إن $k_1 > 0\ ,\ k_2 > 0$

المعلمتان	الشكل
$k_1 < 1$ و $k_2 > 1$	شبه أسي (ينحرف نحو اليمين)
$k_1 > 1$ و $k_2 < 1$	شبه أسي (ينحرف نحو اليسار)
$k_1 = 1$ و $k_2 < 1$	انحدار نحو الأسفل (ينحرف نحو اليمين)
$k_1 < 1$ و $k_2 = 1$	صعود نحو الأعلى (ينحرف نحو اليسار)
$k_1 < 1$ و $k_2 < 1$	شكل حوض الاستحمام
$k_1 > 1$ و $k_2 > 1$	شبه جاما (ينحرف نحو اليسار واليمين)
$k_1 > 1$ و $k_2 > 1$ و $k_1 = k_2$	متماثل، شبه طبيعي
$k_1 = k_2 = 1$	منتظم

يقتصر النقاش في هذا الفصل عندما يكون $k_1 > 1$ و $k_2 > 1$.

المثال (٦-١). على افتراض أن توزيع بيتا القياسي ذو المعلمتين $k_1 = 1.4$ و $k_2 = 2.0$ حيث يتراوح المتغير العشوائي ما بين 0 و1. يتم حساب متوسط ومنوال هذا التوزيع على النحو التالي:

$$\mu = 1.4/3.4 = 0.412$$
$$\tilde{\mu} = 0.4/1.4 = 0.286$$

وتكون دالة الكثافة الاحتمالية على النحو التالي:

$$f(x) = x^{0.4}\,(1-x)^{1.0}\Big/B(1.4,2.0)$$

حيث يتم تحقيق دالة بيتا كما هو مبين أدناه:

$$\begin{aligned} B(1.4,2.0) &= \Gamma(1.4)\Gamma(2.0)\Big/\Gamma(3.4) \\ &= 0.887\times 1.000\Big/2.981 \\ &= 0.298 \end{aligned}$$

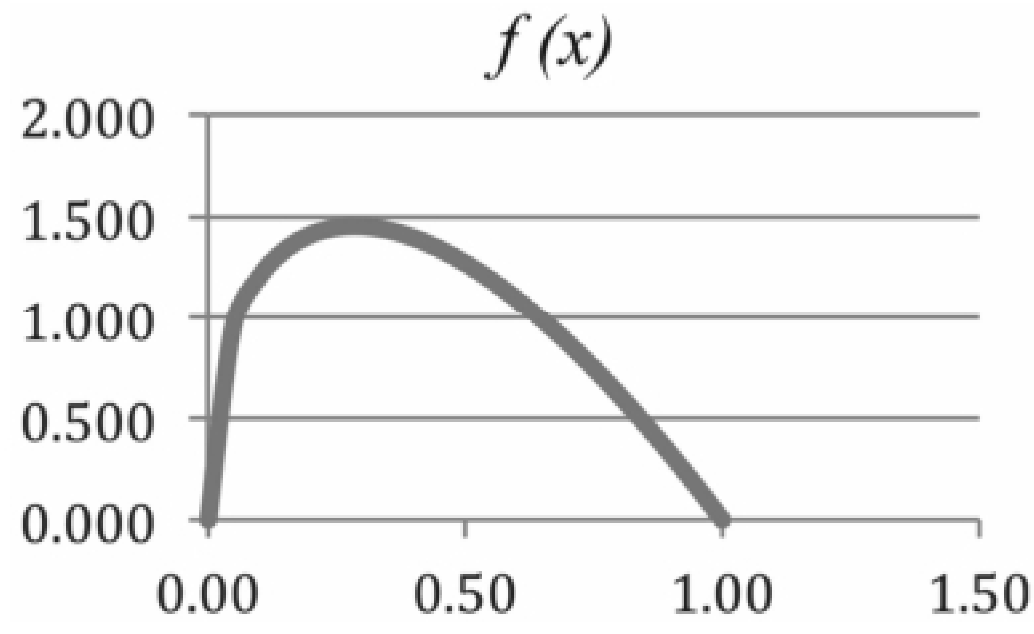
الشكل (٦-١) يمثل دالة الكثافة الاحتمالية لتوزيع بيتا القياسي.

المثال (٦-٢). عند تطبيق توزيع بيتا باستخدام المعالم المبينة في المثال (٦-١)، وبافتراض أن الحدين هما $a=10$ و $b=50$، وبالتالي يصبح المتغير:

$$w = 10 + x(50-10)$$

ويكون متوسط ومنوال W على النحو التالي:

$$\begin{aligned} \mu &= 10 + 0.412(40) = 26.46 \\ \tilde{\mu} &= 10 + 0.286(40) = 21.44 \end{aligned}$$



شكل (٦-١). توزيع بيتا القياسي عندما يكون $k_1 = 1.4$ و $k_2 = 2.0$.

يبيّن الجدول التالي دالة الكثافة الاحتمالية $f(w)$، وتقدير الاحتمال التراكمي $F(w)$، بالنسبة للقيم المُختارة للمتغير w. حيث يتم الحصول على الاحتمال التراكمي $F(w)$ باستخدام العلاقة:

$$F(w_0) \approx \sum_{w=10}^{w_0} f(w) \Bigg/ \sum_{w=10}^{50} f(w)$$

حيث يمتد الجمع على جميع دوال الكثافة الاحتمالية المُدرجة من 10 إلى 50. مع ملاحظة أن الكثافة الاحتمالية $f(w)$ مساوية لـ $f(x)$ عندما يكون $w = (10 + 40x)$.

w	$f(w)$	$\sum f(w)$	$F(w)$
10	0.000	0	0.000
12	0.963	0.963	0.049
14	1.204	2.167	0.110
16	1.337	3.505	0.178
18	1.412	4.917	0.249
20	1.448	6.364	0.322
22	1.453	7.818	0.396
24	1.435	9.253	0.469
26	1.398	10.650	0.540
28	1.343	11.993	0.608
30	1.273	13.267	0.672
32	1.191	14.457	0.732
34	1.096	15.553	0.788
36	0.990	16.543	0.838
38	0.874	17.417	0.882
40	0.749	18.166	0.920
42	0.615	18.781	0.951
44	0.472	19.253	0.975
46	0.322	19.575	0.992
48	0.165	19.740	1.000
50	0.000	19.740	1.000

باستخدام النتائج المبيّنة في الجدول، يلاحظ أنه على سبيل المثال النقطة 0.95% للمتغير w هي تقريباً عندما يكون $w = 42.0$ وبالتالي

$$P(W \leq 42.0) \approx 0.95$$

ويمكن الحصول على تقدير أكثر دقة للدالة $F(w)$ من خلال تطبيق مزيدٍ من قيم $w, f(w)$ و $F(w)$.

٦-٥ بيانات العيّنة

عندما تتوفر بيانات العينة من توزيع بيتا ذي الحدين (a,b) على أنها $(w,\ldots,w_n)$، تكون القياسات التي تم جمعها على النحو التالي:

$\overline{w} =$ متوسط العينة

$\tilde{w} =$ منوال العينة

ويتم التوصل إلى القياسات المقابلة لـتوزيع بيتا القياسي X ذو الحدين (0,1) من خلال:

$$\overline{x} = (\overline{w} - a)\big/(b-a)$$
$$\tilde{x} = (-a)\big/(b-a)$$

٦-٦ تقدير المعالم عند توفر بيانات العيّنة

عند توفر بيانات العيّنة من توزيع بيتا ذي الحدين (a,b)، تكون قياسات العينة اللازمة هي المتوسط $\overline{w}$، والمنوال $\tilde{w}$. وتتحول هذه القياسات إلى متوسط ومنوال بيتا القياسي ($\tilde{x}$ و $\overline{x}$)، ثم يتم إدخالها إلى معادلات متوسط ومنوال المجتمع اللتين تم إدراجهما سابقاً لتشكيل العلاقات المبينة أدناه:

$$\overline{x} = \hat{k}_1\big/(\hat{k}_1 + \hat{k}_2)$$

$$\tilde{x} = \left(\hat{k}_1 - 1\right)\Big/\left(\hat{k}_1 + \hat{k}_2 - 2\right)$$

وباستخدام بعض العمليات الجبرية، يتم استخدام العلاقات المبينة أعلاه لتقدير معلمتي بيتا (k_1, k_2) كما هو مبيّن أدناه:

$$\hat{k}_1 = \overline{x}\left[2\tilde{x} - 1\right]\Big/\left[\tilde{x} - \overline{x}\right]$$
$$\hat{k}_2 = (1-\overline{x})\hat{k}_1\big/\overline{x}$$

المثال (٦-٣). يوجد لدى الباحث بيانات عينة تتبع توزيع بيتا ذي الحدين $(a,b)=(0,100)$ والتي تعطي القياسات التالية:

$$\overline{w} = 30$$
$$\tilde{w} = 20$$

إن تحويل القياسات إلى ما يقابلها في بيتا القياسي يعطي ما يلي:

$$\overline{x} = (30-0)\big/(100-0) = 0.3$$
$$\tilde{x} = (20-0)\big/(100-0) = 0.2$$

ومن ثم يتم حساب معلمتي بيتا على النحو التالي:

$$k_1 = 0.3(2\times 0.2-1)\big/(0.2-0.3) = 1.8$$
$$k_2 = (1-0.3)1.8\big/0.3 = 4.2$$

وأخيراً تكون دالة الكثافة الاحتمالية كما هو مبيّن أدناه:

$$f(x) = x^{0.8}(1-x)^{3.2}\big/B(1.8,4.2)$$

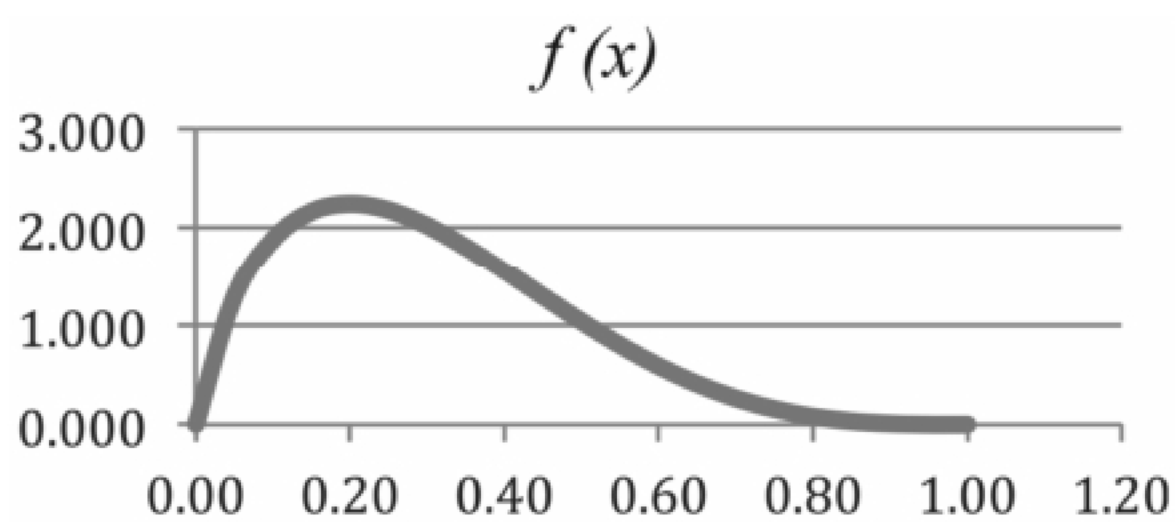
حيث إن

$$\begin{aligned} B(1.8,4.2) &= \Gamma(1.8)\Gamma(4.2)\big/\Gamma(6.0) \\ &= 0.931\times 7.757\big/120.000 \\ &= 0.060 \end{aligned}$$

يبيّن الجدول أدناه قيم دالة الكثافة الاحتمالية المختارة $f(x)$، والاحتمال التراكمي $F(x)$ بالنسبة لقيم x المُختارة. و يتم حساب $F(x)$ كما هو مبيّن في المثال (٦-٢)، كما تم تمثيل الخط البياني لدالة الكثافة الاحتمالية في الشكل (٦-٢).

x	$f(x)$	$\sum f(x)$	$F(x)$
0.00	0.000	0	0.000
0.05	1.283	1.283	0.065
0.10	1.879	3.162	0.160

x	$f(x)$	$\sum f(x)$	$F(x)$
0.15	2.165	5.327	0.269
0.20	2.244	7.571	0.382
0.25	2.182	9.754	0.492
0.30	2.025	11.779	0.594
0.35	1.807	13.586	0.685
0.40	1.556	15.142	0.764
0.45	1.295	16.437	0.829
0.50	1.038	17.475	0.882
0.55	0.800	18.275	0.922
0.60	0.588	18.863	0.952
0.65	0.409	19.272	0.972
0.70	0.265	19.537	0.986
0.75	0.156	19.693	0.994
0.80	0.081	19.774	0.998
0.85	0.034	19.807	0.999
0.90	0.010	19.817	1.000
0.95	0.001	19.818	1.000
1.00	0.000	19.818	1.000



شكل (٦-٢). توزيع بيتا عندما يكون $k_1 = 1.8$ و$k_2 = 4.2$.

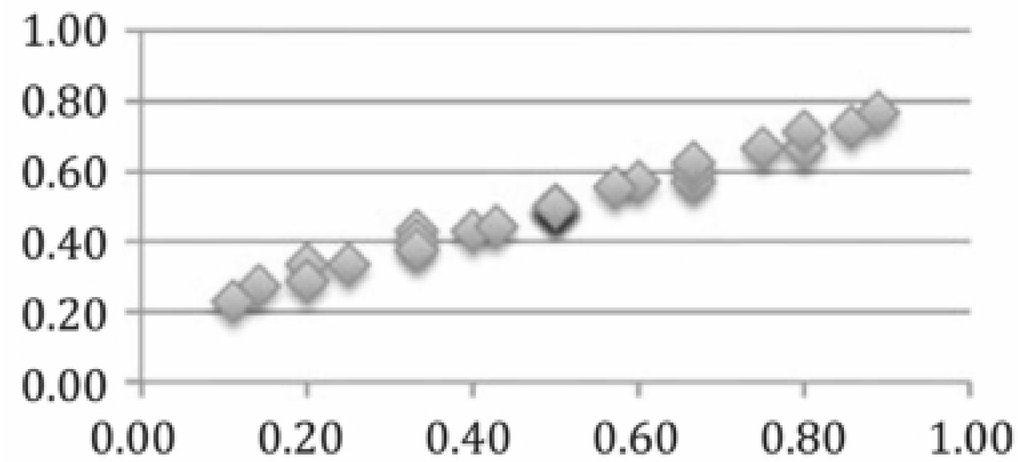
٦-٧ تقدير انحدار المتوسط من المنوال

يبيّن الجدول أدناه تشكيلة لمعلمتي بيتا (k_1, k_2) من مدى القيم المثلى. كما يبيّن القيم المحسوبة للمنوال $\tilde{\mu}$ والمتوسط μ. باستخدام هذه البيانات نوجد الانحدار الخطي للتنبؤ بقيمة المتوسط من المنوال حيث يمثل الشكل (٦-٣) الخط البياني للبيانات والذي يظهر فيه المنوال

على محور x والمتوسط على محور y بحيث تعطي النتيجة علاقة مترابطة جداً وتكون على النحو التالي:

$$\mu = 0.175 + 0.645\tilde{\mu}$$

k_1	k_2	$\tilde{\mu}$	μ
1.5	1.5	0.50	0.50
2	1.5	0.67	0.57
3	1.5	0.80	0.67
4	1.5	0.86	0.73
5	1.5	0.89	0.77
1.5	2	0.33	0.43
2	2	0.50	0.50
3	2	0.67	0.60
4	2	0.75	0.67
5	2	0.80	0.71
1.5	3	0.20	0.33
2	3	0.33	0.40
3	3	0.50	0.50
4	3	0.60	0.57
5	3	0.67	0.63
1.5	4	0.14	0.27
2	4	0.25	0.33
3	4	0.40	0.43
4	4	0.50	0.50
5	4	0.57	0.56
1.5	5	0.11	0.23
2	5	0.20	0.29
3	5	0.33	0.38
4	5	0.43	0.44
5	5	0.50	0.50



شكل (٦-٣). المنوال على محور x مقابل المتوسط على محور y .

٦-٨ تقدير المعالم عند عدم توفر بيانات

عند عدم توفر بيانات العينة، يطلب المحلل من أحد الخبراء تقديم بعض المعلومات المتعلقة بشكل التوزيع. تتمثل إحدى الطرق في تقدير القيمة المرجّحة (المنوال) والمُشار إليها بـ $\tilde{w}$ ، وكذلك الحدين الأدنى والأعلى (a,b) . باستخدام هذا التقدير يتم التوصل إلى منوال توزيع بيتا من العلاقة أدناه:

$$\tilde{x} = (\tilde{w} - a)\Big/(b - a)$$

وبعد ذلك يطبق المحلل نتيجة الانحدار لتقدير قيمة متوسط x على النحو التالي:

$$\overline{x} = 0.175 + 0.645\,\tilde{x}$$

تتيح هذه الخطوة للمحلل التوصل إلى تقدير المقياسين الرئيسين لتوزيع بيتا القياسي: $\tilde{x}$ و $\overline{x}$ بحيث يصبح المحلل الآن قادراً على حساب تقدير معلمتي بيتا على النحو التالي:

$$\hat{k}_1 = \overline{x}\left[2\tilde{x} - 1\right]\Big/\left[\tilde{x} - \overline{x}\right]$$
$$\hat{k}_2 = \left[1 - \overline{x}\right]\hat{k}_1\Big/\overline{x}$$

المثال (٦-٥). يقوم الباحث بإعداد دراسة ويحتاج إلى استخدام متغير بيتا w إلا أنه لا يوجد لديه بيانات العينة لتقدير المعلمتين. يسعى الباحث للحصول على قيمة w حيث إن الاحتمال التراكمي هو 0.95. ولقد حصل الباحث على التقدير التالي المتعلق بتوزيع بيتا.

$a = 50$ هي القيمة الدنيا

$b = 100$ هي القيمة العليا

$\tilde{w} = 80$ هي القيمة المرجّحة

وبالتالي يصبح المنوال المقابل لبيتا القياسي:

$$\begin{aligned}\tilde{x} &= (80-50)\big/(100-50)\\ &= 0.60\end{aligned}$$

ومن ثم باستخدام الانحدار، يتم حساب تقدير متوسط بيتا القياسي على النحو التالي:

$$\begin{aligned}\overline{x} &= 0.175 + 0.645 \times 0.60\\ &= 0.562\end{aligned}$$

من خلال هذه القياسات يتم التوصل إلى تقدير معلمتي بيتا كما هو مبين أدناه:

$$\begin{aligned}k_1 &= 0.562\big[2\times 0.60-1\big]\big/\big[0.60-0.562\big]\\ &= 2.958\\ k_2 &= \big[1-0.562\big]2.958\big/0.562\\ &= 2.305\end{aligned}$$

وأخيراً تكون دالة الكثافة الاحتمالية لـ x على النحو التالي:

$$f(x) = x^{1.958}(1-x)^{1.305}\big/B(2.958, 2.305)$$

حيث إن

$$\begin{aligned}B(2.958, 2.305) &= \Gamma(2.958)\Gamma(2.305)\big/\Gamma(5.263)\\ &= 1.924\times 1.1702\big/35.933\\ &= 0.063\end{aligned}$$

يبيّن الجدول (٦-١) الكثافة الاحتمالية $f(x)$، والاحتمال التراكمي $F(x)$، بالنسبة لقيم x المختارة بحيث يمثل الشكل (٦-٤) دالة الكثافة الاحتمالية للمتغير X. يتم التوصل إلى النقطة 0.95% للمتغير X كما هو مبيّن أدناه:

$$\begin{aligned} x_{0.95} &= 0.85 + 0.05(0.950 - 0.934)\Big/(0.973 - 0.934) \\ &= 0.871 \end{aligned}$$

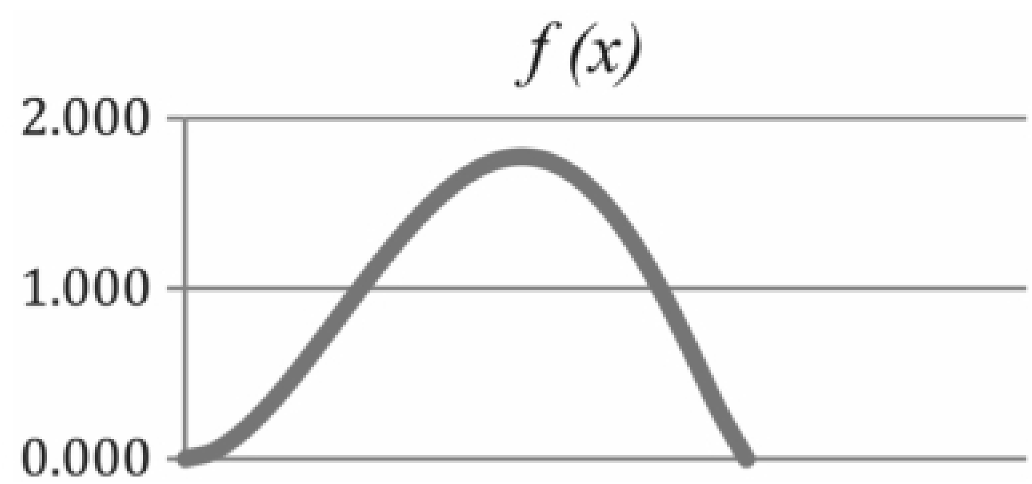
وتكون النقطة 0.95% المقابلة بالنسبة w هي:

$$\begin{aligned} w_{0.95} &= 50 + 0.871\Big/(100 - 50) \\ &= 93.55 \end{aligned}$$

وبالتالي، $P(w \leq 93.55) = 0.95$

جدول (٦-١). قائمة توزيع بيتا القياسي x، مع الكثافة الاحتمالية $f(x)$ والاحتمال التراكمي $F(x)$.

x	$f(x)$	$\sum f(x)$	$F(x)$
0.00	0.00	0.00	0.00
0.05	0.04	0.04	0.00
0.10	0.15	0.20	0.01
0.15	0.31	0.51	0.03
0.20	0.51	1.02	0.05
0.25	0.73	1.75	0.09
0.30	0.95	2.69	0.13
0.35	1.16	3.86	0.19
0.40	1.36	5.22	0.26
0.45	1.53	6.75	0.34
0.50	1.66	8.41	0.42
0.55	1.75	10.16	0.51
0.60	1.77	11.93	0.60
0.65	1.74	13.68	0.68
0.70	1.65	15.33	0.77
0.75	1.49	16.82	0.84
0.80	1.26	18.08	0.90
0.85	0.98	19.05	0.95
0.90	0.64	19.70	0.99
0.95	0.29	19.99	1.00
1.00	0.00	19.99	1.00



شكل (٦-٤). الكثافة الاحتمالية لتوزيع بيتا عندما يكون $k_1 = 2.958$ ، $k_2 = 2.305$ ، و $a = 50$ و$b = 100$.

٦-٩ ملخص

هناك عدة أشكال لتوزيع بيتا. يركّز هذا الفصل على الأشكال التي تنحدر نحو اليسار واليمين فقط. إن لمتغير بيتا مدى يتراوح ما بين a وb، كما أن لمتغيره المقابل بيتا القياسي مدى يتراوح ما بين 0 و1. تكون الحسابات الرئيسية من خلال بيتا القياسي ويمكن تحويلها بسهولة إلى متغير بيتا. ويتم التوصل إلى قيم المعالم من خلال متوسط ومنوال التوزيع، وهو ما يتيح تحديد الكثافة الاحتمالية، والتي تتضمن أيضاً دالة بيتا. تتألف دالة بيتا من دالة جاما التي لا يمكن حساب قيمتها بسهولة إلا من خلال خوارزمية الكمبيوتر. ولا يوجد في الاحتمال التراكمي لبيتا حل محدد الشكل، وهناك ضرورة إلى الطرق الكمّية لإجراء الحسابات، وعند توفر بيانات العينة، يتم استخدام المتوسط والمنوال لتقدير قيم المعالم، وعند عدم توفر بيانات العينة، يقوم أحد الأشخاص المطلعين بتقديم الحدين (a، b) والقيمة المرجّحة. ومن خلال هذه البيانات، يتم تقدير المعالم المجهولة.

الفصل السابع

توزيع ويبل
WEIBULL DISTRIBUTION

٧-١ مقدمة

تمَّ تقديم توزيع ويبل رسمياً عام 1939 من قبل عالم الرياضيات السويدي والودي ويبل (Waloddi Weibull). وقد استخدم الفرنسي موريس فريشيت (Maurice Frechet) التوزيع في عام 1927، وقام كلٌّ من روسين و راملر (R. Rosin and E. Rammler) بتطبيقه في عام 1933. هناك صيغتان لتوزيع ويبل تتراوحان ما بين شبه أسي إلى شبه طبيعي، ويأخذ المتغير العشوائي W قيم γ أو أكبر. إن لتوزيع ويبل القياسي ذي المتغير X قيم صفر أو أكبر. كما أنَّ لكلا التوزيعين المعلمتين ذاتها وهما (k_1, k_2) بحيث يشكلان شكل التوزيع. وعندما يكون $k_1 \leq 1$، يكون منوال توزيع ويبل القياسي صفراً، ويكون الشكل شبه أسي. بينما عندما يكون $k_1 > 1$، يكون المنوال أكبر من الصفر، وعندما يكون $k_1 = 3$ أو أكبر، يكون الشكل شبه طبيعي. وسوف يتم توضيح المعادلات الرياضية المتعلقة بدالة الكثافة الاحتمالية ودالة الاحتمال التراكمي بسهولة. إلا أنه ليس من السهل حساب متوسط وتباين التوزيع ويتطلب ذلك استخدام دالة جاما. كما يبيّن هذا الفصل طرق تقدير المعالم γ, k_1, k_2 عند توفر بيانات العينة. وعند عدم توفر البيانات، يمكن تقديم أحد الخبراء تقريبات لبعض قياسات التوزيع وسيتم توضيح الطرق المتعلقة بكيفية تقدير قيم هذه المعالم.

٧-٢ أساسيات

تم إدراج المتغير العشوائي لتوزيع ويبل على أنه W بحيث يكون مداه (γ وما فوق)، في حين أن مدى المتغير العشوائي المقابل لتوزيع ويبل القياسي X ، يتراوح ما بين (0 وما فوق). وتكون معلمات توزيع ويبل على النحو التالي:

معالم التوزيع $= (k_1 > 0, k_2 > 0)$

$\gamma =$ معلمة المكان الأدنى للمتغير W

$0 =$ معلمة المكان الأدنى للمتغير X

وتكون العلاقة بين W و X على النحو التالي:

$$x = w - \gamma$$
$$w = x + \gamma$$

٣-٧ توزيع ويبل القياسي

إن دالة الكثافة الاحتمالية لتوزيع ويبل القياسي تكون على النحو التالي:

$$f(x) = k_1 k_2^{-k_1} x^{k_1 - 1} \exp\left[-(x/k_2)^{k_1}\right] \qquad x > 0$$

وبالتالي فإن دالة التوزيع التراكمية:

$$F(x) = 1 - \exp\left[-(x/k_2)^{k_1}\right] \qquad x > 0$$

كما يعرف متوسط وتباين المتغير العشوائي X على النحو التالي:

$$\mu = k_2/k_1\, \Gamma(1/k_1)$$
$$\sigma^2 = k_2^2/k_1\left[2\Gamma(2/k_1) - 1/k_1 \Gamma(1/k_1)^2\right]$$

ويكون المنوال على الشكل التالي:

$$\begin{aligned} \tilde{\mu} &= k_2\left[(k_1 - 1)/k_1\right]^{1/k_1} && k_1 \geq 1 \\ &= 0 && k_1 < 1 \end{aligned}$$

يُشار إلى النقطة المئوية α بالنسبة للمتغير X بالمقدار $x\alpha$ ، ويتم التوصل إليها على النحو المبيّن أدناه:

$$F(x\alpha) = \alpha = 1 - \exp\left[-(x\alpha/k_2)^{k_1}\right]$$

وعند تطبيق بعض العمليات الجبرية،

$$x\alpha = -k_2 \ln(1-\alpha)^{1/k_1}$$

حيث إن $\ln$ هو اللوغاريتم الطبيعي.

المثال (٧-١). يستخدم المحلل توزيع ويبل من خلال المعالم التالية:

$$\begin{aligned} \gamma &= 100 \\ k_1 &= 1.97 \\ k_2 &= 28.66 \end{aligned}$$

يصبح توزيع ويبل القياسي ذي الصلة:

$$x = w - 100,$$

ودالة الكثافة الاحتمالية للمتغير X هي:

$$f(x) = 1.97 \times 28.66^{1.97} x^{0.97} \exp\left[-(x/28.66)^{1.97}\right] \qquad x > 0$$

كما يصبح الاحتمال التراكمي لما هو أقل أو يساوي $x = 50$ بالصورة:

$$\begin{aligned} F(50) &= 1 - \exp\left[-(50/28.66)^{1.97}\right] \\ &= 0.95 \end{aligned}$$

وبالتالي فإن احتمال أن يكون W أقل من 150 هو:

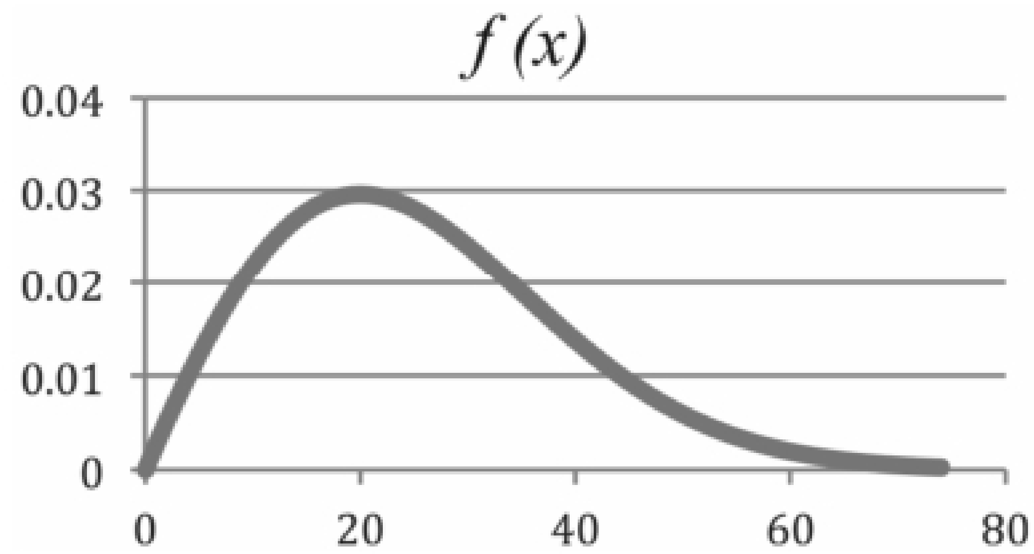
$$P(w \le 150) = 0.95$$

يتم حساب النقطة 0.95% للمتغير X كما هو مبين أدناه:

$$x_{0.95} = -28.66 \times \ln(1-0.95)^{1/1.97} = 50$$

وتكون القيمة المقابلة للمتغير العشوائي W كما يلي:

$$w_{0.95} = 100 + 50 = 150$$



شكل (٧-١). دالة الكثافة الاحتمالية لـلمتغير العشوائي X عندما $k_1 = 1.97$ و $k_2 = 28.66$.

٧-٤ بيانات العيّنة

عند توفر بيانات العيّنة $(w_1,\dots,w_n)$ ، تكون قياسات ويبل ذات الصلة على النحو التالي:

$\gamma =$ معلمة موقع w

$\tilde{w} =$ منوال w

$w_{0.5} =$ وسيط w (النقطة المئوية 0.5)

٧-٥ تقدير معلمة γ عند توفر بيانات العيّنة

يقدّم زاناكيز (1979) ، [Zanakis (1979) p 101–116] طريقة لتقدير معلمة الموقع (الدنيا) للمتغير w كما هو مبيّن أدناه:

$$\gamma = \left[w(1)w(n) - w(k)^2\right] \Big/ \left[w(1) + w(n) - 2w(k)\right]$$

حيث إن

$w(1) = \min(w_1,\dots,w_n)$ القيمة الدنيا

$w(n) = \max(w_1,\dots,w_n)$ القيمة العليا

k العدد الصحيح الأصغر حيث إن $w(k) > w(1)$

المثال (٧-٢). يوجد لدى الباحث مشاهدات العيّنة $n = 10$ ويريد استخدام توزيع ويبل في دراسة بحثية بحيث كانت المشاهدات على النحو التالي:

49,37,63,84,71,55,48,62,58,43

إن التركيز الأول يكون على تقدير معلمة الموقع الدنيا γ . وعند تطبيق صيغة زاناكيز نجد أن:

$$w(1) = 37$$
$$w(10) = 84$$
$$w(k) = 43$$

وبالتالي يتم تقدير معلمة الموقع الدنيا على النحو التالي:

$$\gamma = (37 \times 84 - 43^2)\Big/(37 + 84 - 2 \times 43) = 35.97$$

٧-٦ تقدير المعلمتين (k_1, k_2) عند توفر بيانات العيّنة

من الممكن تقدير المعلمتين (k_1, k_2) عندما يكون $k_1 \geq 1$، كذلك عندما يكون المنوال أكبر من الصفر. وعليه يتم تحويل مقياسي عيّنة ويبل $(\tilde{w}, w_{0.5})$ إلى مقاييس توزيع ويبل القياسي كما هو مبين أدناه:

$$X \text{ منوال } = \tilde{x} = \tilde{w} - \gamma$$

$$X \text{ وسيط } = x_{0.5} = w_{0.5} - \gamma \quad (\text{النقطة } 0.50\% \text{ للمتغير } X)$$

عندما يكون $k_1 < 1$، فإن منوال X يكون عندما $x = 0$. يتمثل التحليل هنا في أنه عندما يكون $k_1 \geq 1$ ومنوال X أكبر من الصفر، يتم قياس المنوال على النحو المبين أدناه:

$$\tilde{x} = k_2 \left[(k_1 - 1)\big/k_1\right]^{1/k_1}$$

وباستخدام بعض العمليات الجبرية تصبح k_2

$$k_2 = \tilde{x}\Big/\left[(k_1 - 1)\big/k_1\right]^{1/k_1}$$

إن النقطة المئوية α بالنسبة للمتغير X هي x_α، ويتم تحقيقها من خلال ما يلي:

$$F(x_\alpha) = 1 - \exp\left[-(x_\alpha\big/k_2)^{k_1}\right] = \alpha$$

وبالتالي

$$\ln(1 - \alpha) = -(x_\alpha\big/k_2)^{k_1}$$

وعند تطبيق مزيد من العمليات الجبرية والحل بالنسبة لـ k_2 نجد أن:

$$k_2 = x_\alpha \Big/ Ln\left[1\Big/\left(1-\alpha\right)\right]^{\frac{1}{k_1}}$$

نلاحظ أن المجهول الوحيد في المعادلة أدناه هو k_1

$$\tilde{x}\Big/\left[(k_1-1)\Big/k_1\right]^{1/k_1} = x_\alpha \Big/ Ln[1\Big/(1-\alpha)]^{1/k_1}$$

حيث إن

$$\tilde{x}\Big/x_\alpha = \left\{(k_1-1)\Big/[k_1 \times \ln[1\Big/(1-\alpha)]]\right\}^{1/k_1}$$

عند استبدال $\alpha = 0.50$ و $x\alpha = x_{0.5}$ ينتج ما يلي:

$$\tilde{x}\Big/x_{0.5} = \left\{(k_1-1)\Big/\left[k_1 \times \ln(2)\right]\right\}^{1/k_1}$$

الحل من أجل k_1: تم إجراء بحث مكرّر لإيجاد قيمة k_1 حيث إن الجانب الأيمن من المعادلة أعلاه يساوي الجانب الأيسر بحيث يكون $\hat{k}_1$

الحل من أجل k_2 بعد إيجاد $\hat{k}_1$، يتم الحصول على المعلمة الأخرى k_2 من خلال العلاقة:

$$\hat{k}_2 = \tilde{x}\Big/[(\hat{k}_1-1)\Big/\hat{k}_1]^{1/\hat{k}_1}$$

المثال (٧-٣). باستخدام بيانات المثال (٧-٢)، يحتاج الباحث بعد ذلك إلى تقدير معلمتي ويبل (k_1, k_2)، ويتطلب ذلك تقدير المنوال $\tilde{x}$ والوسيط $x_{0.5}$ من توزيع ويبل القياسي.

تتمثل إحدى طرق تقدير المنوال في تنظيم البيانات في ست مجموعات فرعية بحجم 10 كما هو مبين أدناه ومن ثم حساب عدد أرقام المشاهدات في كل مجموعة فرعية بحيث تكون المجموعة الفرعية ذات الرقم الأكبر هي موقع المنوال. في حين يمثل متوسط المشاهدات في المجموعة الفرعية المُختارة إحدى طرق تقدير قيمة المنوال. في المثال السابق، تكون حسابات تقدير توزيع ويبل وتوزيع ويبل القياسي على النحو المبين أدناه:

$$\tilde{w} = (48 + 49 + 43)/3 = 46.67$$
$$\tilde{x} = (46.67 - 35.97) = 10.70$$

المجموعة الفرعية	$80-89$	$70-79$	$60-69$	$50-59$	$40-49$	$30-29$
الأرقام	1	1	2	2	3	1

إن الخطوة التالية هي تقدير وسيط البيانات وهي النقطة 0.50% المُشار إليها بـ$x_{0.50}$. وللحصول على ذلك، يتم ترتيب نقاط البيانات $n = 10$ من الأدنى إلى الأعلى كما هو مبيّن أدناه:

37,43,48,49,55,58,62,63,71,84

إن القيمة المتوسطة هي ما بين المشاهدة الخامسة والسادسة. ويكون تقدير وسيط توزيع ويبل ووسيط توزيع ويبل القياسي على النحو التالي:

$$w_{0.5} = (55 + 58)/2 = 56.50$$
$$x_{0.5} = (56.50 - 35.97) = 20.53$$

من خلال مقياسي ويبل القياسي $(\tilde{x}, x_{0.5})$ المأخوذين من البيانات، أصبح من الممكن حالياً تقدير قيم معلمتي توزيع ويبل k_1, k_2 باستخدام الطريقة المكرّرة المبيّنة سابقاً. تصبح نسبة المنوال إلى الوسيط:

$$\tilde{x}/x_{0.5} = 10.70/20.53 = 0.52$$

وباستخدام المعادلة المبيّنة سابقاً نجد أن:

$$\tilde{x}/x_{\alpha} = \left\{(k_1 - 1)/[k_1 \times \ln[1/(1-\alpha)]]\right\}^{1/k_1}$$

باعتبار أن الجانب الأيسر من المعادلة يساوي 0.52، فقد تم اجراء بحث مكرّر عن الجانب الأيمن باعتبار أن $\alpha = 0.50$، وقيم k_1 من 1.1 وما فوق. عندما يكون الجانب الأيمن قريب من 0.52، يتم تحديد قيمة k_1 في الجدول أدناه ويحدث ذلك عندما يكون $\hat{k}_1 = 1.55$.

الجانب الأيمن	k_1
0.128	1.1
0.233	1.2
0.325	1.3
0.408	1.4
0.487	1.5
0.562	1.6
0.636	1.7
0.707	1.8
0.779	1.9
0.849	2.0

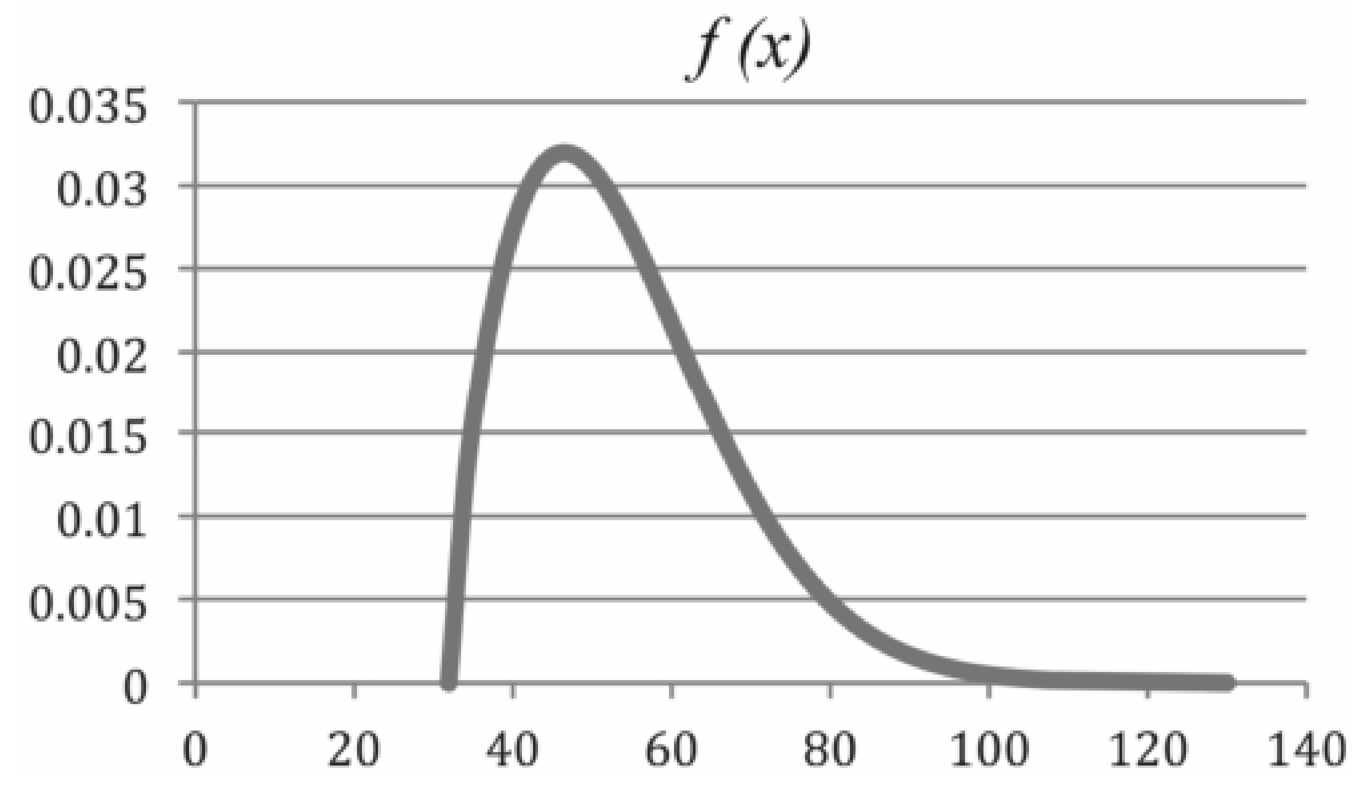
عند تقدير k_1 ، أصبح من الممكن الآن تقدير k_2 من خلال المعادلة:

$$\begin{aligned}\hat{k}_2 &= \tilde{x}\Big/\left[(\hat{k}_1 - 1)\Big/\hat{k}_1\right]^{1/\hat{k}_1} \\ &= 10.70\Big/\left[0.55\Big/1.55\right]^{1/1.55} \\ &= 20.87\end{aligned}$$

وبالتالي تصبح معالم توزيع ويبل التي تم تقديرها من بيانات العيّنة على النحو التالي:

$$(\hat{\gamma}, \hat{k}_1, \hat{k}_2) = (35.97, 1.55, 20.87)$$

يمثّل الشكل (٧-٢) الرسم البياني لتوزيع ويبل ذي المعالم السابقة.



شكل (٧-٢). توزيع ويبل عندما يكون $\gamma = 35.97$ ، $k_1 = 1.55$ ، و $k_2 = 20.87$.

٧-٧ تقدير المعالم عند عدم توفر بيانات

عندما لا تتوفر بيانات العيّنة ويريد الباحث تطبيق توزيع ويبل، يكون هناك حاجة إلى قياسات شكل التوزيع ويتم توفير أفضل تقدير على النحو التالي:

$\gamma =$ معلمة موقع w

$\tilde{w} =$ القيمة المرجّحة للمتغير w (المنوال)

$(w\alpha, \alpha) =$ النقطة المئوية α لـ w

باستخدام هذه القياسات، يكون التحويل إلى توزيع ويبل القياسي على النحو التالي:

$\tilde{x} =$ تقدير منوال X

$(x\alpha, \alpha) =$ النقطة المئوية α للمتغير X

باستخدام $(\tilde{x}, x\alpha, \alpha)$ يمكن تطبيق الخوارزمية المكرّرة المدرجة أعلاه لتقدير معلمي ويبل (k_1, k_2).

المثال (٧-٤). على افتراض أن المحلل يسعى لاستخدام توزيع ويبل ولا توجد بيانات العيّنة لتقدير المعلمتين. لذا فإنه بمساعدة أحد الخبراء تم الحصول على التقدير التالي:

$\gamma =$ القيمة الدنيا $= 10$

$\tilde{w} =$ القيمة المرجّحة $= 20$

$w_{0.9} =$ النقطة $0.90\% = 30$

$\alpha = 0.90$

باستخدام التقدير أعلاه يصبح منوال توزيع ويبل القياسي:

$$\tilde{x} = (20 - 10) = 10$$

في المعادلة المبيّنة أدناه، الجانب الأيسر يساوي $10/20 = 0.50$، بينما يتم البحث عن الجانب الأيمن مع $\alpha = 0.90$ من أجل قيمة k_1 (من 1.1 إلى 2.1) وعندما يكون الجانب الأيمن قريباً من 0.50 نحصل على $k_1 \approx 2.06$.

$$\tilde{x}/x_\alpha = \left\{(k_1 - 1)/[k_1 \times \ln[1/(1-\alpha)]\right\}^{1/k_1}$$

الجانب الأيمن	k_1
0.042	1.1
0.085	1.2
0.129	1.3
0.173	1.4
0.218	1.5
0.265	1.6
0.313	1.7
0.363	1.8
0.413	1.9
0.466	2.0
0.519	2.1

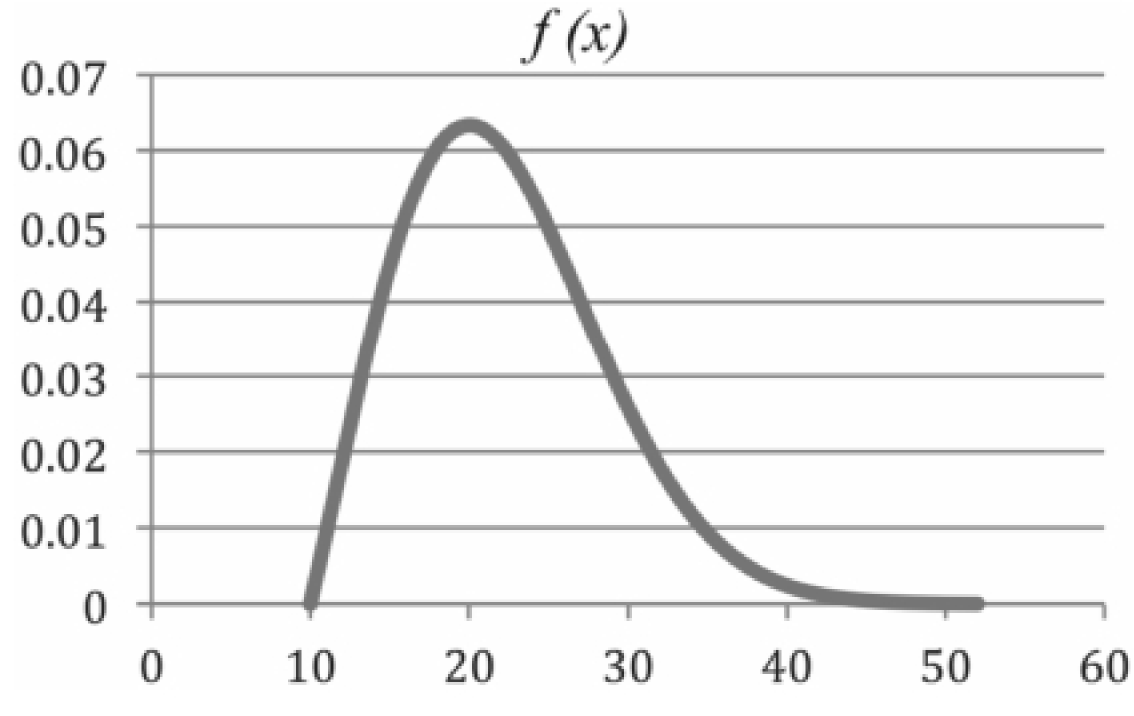
باستخدام $\hat{k}_1 = 2.06$ تعطي المعادلة أدناه تقدير $\hat{k}_2$.

$$\begin{aligned}\hat{k}_2 &= \tilde{x} \Big/ \left[(\hat{k}_1 - 1) \Big/ \hat{k}_1\right]^{1/\hat{k}_1} \\ &= 10 \Big/ [1.06/2.06]^{1/2.06} \\ &= 13.80\end{aligned}$$

وأخيراً تصبح معالم توزيع ويبل على الصورة:

$$(\hat{\gamma}, \hat{k}_1, \hat{k}_2) = (10, 2.06, 13.80)$$

وقد تم توضيح الرسم البياني لتوزيع ويبل في الشكل (٧-٣) باستخدام المعالم السابقة.



شكل (٧-٣). توزيع ويبل عندما يكون $\hat{\gamma} = 10$ **، و** $\hat{k}_1 = 2.06$ **، و** $\hat{k}_2 = 13.80$.

٧-٨ ملخص

يتراوح مدى المتغير العشوائي لتوزيع ويبل w ما بين γ وما فوق، ويكون المتغير العشوائي لتوزيع ويبل القياسي x صفراً أو أكبر. ولكل من التوزيعين المعلمتين ذاتها (k_1, k_2) واللتين تشكلان التوزيع. غالباً ما تكون التعاملات الرياضية مع توزيع ويبل القياسي. وعندما يكون k_1 أقل من واحد أو مساوياً له، يكون شكل توزيع ويبل القياسي شبه أسي. وعندما يكون k_1 أكبر من واحد، يكون المنوال أكبر من صفر. وعندما يصبح k_1 ثلاثة أو أكبر، يبدو الشكل شبه طبيعي. لقد بيّن هذا الفصل الطريقة المتعلقة بكيفية استخدام بيانات العيّنة لتقدير معلمة الموقع γ بالنسبة لتوزيع ويبل. إضافةً إلى ذلك عند توفر تقدير المنوال والنقطة المئوية α، يتم استخدام خوارزمية مكرّرة لتقدير المعلمتين (k_1, k_2). كما قدّم الفصل أمثلة عن كيفية استخدام التوزيع عند توفر بيانات العيّنة وكذلك عند عدم توفرها.

الفصل الثامن

التوزيع الطبيعي
NORMAL DISTRIBUTION

٨-١ مقدمة

في عام 1809 قدم كارل فريدريك غاوس (Carl Friedrich Gauss) طريقة المربعات الصغرى، وطريقة مقدّر الإمكان الأعظم، والتوزيع الطبيعي، وغالباً ما يُشار إليها بتوزيع غاوس. إن التوزيع الطبيعي هو التوزيع الأكثر استخداماً في جميع التخصصات. كما أن للتوزيع الطبيعي المتغير العشوائي X ذي المعالم μ و σ حيث تمثلان المتوسط و الانحراف المعياري على الترتيب. في حين أن للتوزيع الطبيعي القياسي المتغير العشوائي Z ويكون متوسطه صفراً وانحرافه المعياري واحداً. يبيّن الفصل طريقة سهلة للتحويل من X إلى Z ومن Z إلى X، حيث إن الجداول المتعلقة بالتوزيع الطبيعي القياسي متاحة في جميع كتب الإحصاء تقريباً. ولا توجد صيغة محددة الشكل بالنسبة لدالة الاحتمال التراكمية $F(z)$، لذا تم وضع العديد من الطرق الكمّية على مرّ السنين لتقدير هذه الدالة. يقدم هذا الفصل صيغة تقريب هاستينج لإيجاد الاحتمال التراكمي $F(z)$ من z، وهناك صيغة أخرى لتقريب هاستينج لإيجاد z من $F(z)$. عند توفر بيانات العيّنة، يتم استخدام متوسط العينة والانحراف المعياري للعينة لتقدير متوسط التوزيع الطبيعي والانحراف المعياري لها. وعند عدم توفر البيانات، يتم توضيح الطريقة المتعلقة بكيفية تقدير متوسط التوزيع الطبيعي والانحراف المعياري لها من بعض القياسات التقريبية للتوزيع.

٨-٢ أساسيات

يُشار إلى المتغير العشوائي للتوزيع الطبيعي بأنه X ويتراوح مداه ما بين $-\infty$ و$+\infty$. وتكون دالة الكثافة الاحتمالية له على النحو التالي:

$$f(x) = \sqrt{\left(\frac{1}{2\pi\sigma^2}\right)} e^{-0.5\left[(x-\mu)/\sigma\right]^2}$$

حيث إن

$\mu =$ المتوسط

$\sigma^2 =$ التباين

$\sigma =$ الانحراف المعياري

ويكون ترميز التوزيع الطبيعي على النحو المبين أدناه:

$$X \sim N(\mu, \sigma^2)$$

يُشار إلى الاحتمال التراكمي للمتغير العشوائي الذي هو أقل من X أو يساويه بـ $F(x)$، وعندما يكون $x = x_0$ فإن الاحتمال يعرف على النحو التالي:

$$F(x_0) = P(x \le x_0) = \int_{-\infty}^{x_0} f(x)\, dx$$

نشير إلى أنه لا توجد صيغة محددة لما ذُكر أعلاه، وقد تم وضع الطرق الكمّية لإنشاء جداول عن $F(x)$

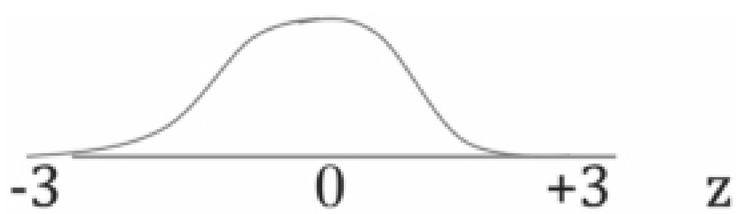
٨-٣ التوزيع الطبيعي القياسي

يُشار إلى المتغير العشوائي للتوزيع الطبيعي القياسي بأنه Z ويكون متوسطه صفراً وانحرافه المعياري واحداً. يمثّل الشكل (٨-١) شكل التوزيع الطبيعي القياسي. وقد تم ترميز المتغير العشوائي Z أدناه:

$$Z \sim N(0,1)$$

كما تم توضيح طريقة التحويل من X (التوزيع الطبيعي) إلى Z (التوزيع الطبيعي القياسي)، والعكس فيما يلي:

$$z = (x - \mu)/\sigma$$
$$x = \mu + z\sigma$$



شكل (٨-١). التوزيع الطبيعي القياسي.

يتم التوصل إلى دالة الكثافة الاحتمالية، ودالة الاحتمال التراكمية، والاحتمال المكمّل كما هو مبين أدناه حيث إن k تمثل قيمة معيّنة للمتغير z .

دالة الكثافة الاحتمالية للمتغير Z = $f(z) = \left(1/\sqrt{2\pi}\right)e^{-z^2/2}$

عندما $z = k$ فإن الاحتمال التراكمي $F(k) = F(z \le k) = \int_{-\infty}^{k} f(z)\,dz$

وعندما $z = k$ فإن الاحتمال المكمّل ـ $H(k) = P(z > k) = 1 - F(k)$

ومن خصائص التوزيع الطبيعي القياسي ما يلي:

$$\int_{-\infty}^{\infty} zf(z)\,dz = 0$$
$$\int_{-\infty}^{\infty} z^2 f(z)\,dz = 1$$

نلاحظ أنه باعتبار أن Z متغير متصل، فإن $F(k) = P(z \le k) = P(z < k)$.

٨-٤ تقريبات هاستينج (Hastings Approximations)

باعتبار أنه لا توجد صيغة محددة الشكل للتوزيع التراكمي $F(z)$ ، فقد تم وضع العديد من الطرق الكمّية لتقدير الاحتمال التراكمي للتوزيع الطبيعي القياسي على مرّ السنين و لقد قدم هاستينج، أبراموويتز وستيجن عام 1964 [Abramowitz and Stegun (1964). P931-936] طريقتين تركزان على العلاقة بين $F(z)$ و Z حيث تم توضيحهما في هذا الفصل.

تقريب $F(z)$ من Z بالنسبة لقيمة محددة Z ، ولإيجاد $F(z)$ يتم استخدام منهج هاستينج التالي:

١.

$$d_1 = 0.0498673470$$
$$d_2 = 0.0211410061$$
$$d_3 = 0.0032776263$$
$$d_4 = 0.0000380036$$
$$d_5 = 0.0000488906$$
$$d_6 = 0.0000053830$$

٢. إذا كان $z \geq 0 : w = z$

وإذا كان $z < 0 : w = -z$

٣. $F = 1 - 0.5\left[1 + d_1 w + d_2 w^2 + d_3 w^3 + d_4 w^4 + d_5 w^5 + d_6 w^6\right]^{-16}$

٤. إذا كان $z \geq 0 : F(z) = F$

وإذا كان $z < 0 : F(z) = 1 - F$

نكرر الخطوات السابقة لـ $F(z)$.

تقريب Z من $F(z)$

هناك تقريب آخر مفيد من هاستينج يوفر منهجاً يعطي المتغير العشوائي Z من قيمة $F(z)$ كما يلي:

١.

$$c_0 = 2.515517$$
$$c_1 = 0.802853$$
$$c_2 = 0.010328$$
$$d_1 = 1.432788$$
$$d_2 = 0.189269$$
$$d_3 = 0.001308$$

.٢

$$H(z) = 1 - F(z)$$
$$H(z) \le 0.5 : H = H(z)$$
$$H(z) > 0.5 : H = 1 - H(z)$$

٣. $t = \sqrt{\ln(1/H^2)}$ ، حيث $\ln$ = لوغاريتم طبيعي.

٤. $w = t - \left[c_0 + c_1 t + c_2 t^2\right] / \left[1 + d_1 t + d_2 t^2 + d_3 t^3\right]$

٥. إذا كان $H(z) \le 0.5 : z = w$

وإذا كان $H(z) > 0.5 : z = -w$

نكرر الخطوات السابقة بالنسبة إلى Z .

٨-٥ جداول التوزيع الطبيعي القياسي

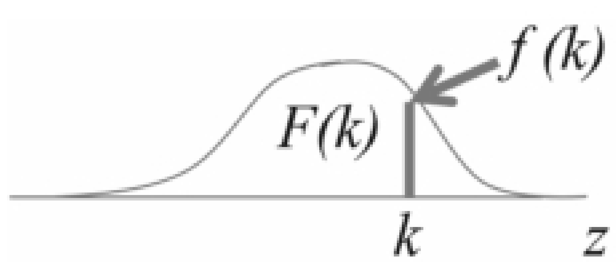
يبيّن الجدول (٨-١) قيم k و $F(k)$ و $H(k)$ و $f(k)$ ، من التوزيع الطبيعي القياسي التي تقع في المدى: $[-0.3,(0.1),+0.3]$. و يمثل الشكل (٨-٢) $F(k)$ ، و $f(k)$ ، و k بالنسبة للمتغير الطبيعي القياسي Z .

جدول (٨-١). الإحصاءات الطبيعية القياسية المرتّبة حسب k ، مع التوزيع التراكمي $F(k)$ ، والاحتمال المكمّل $H(k)$ ، والكثافة الاحتمالية $f(k)$.

k	$F(k)$	$H(k)$	$f(k)$	k	$F(k)$	$H(k)$	$f(k)$
-3.0	0.001	0.999	0.004	0.0	0.500	0.500	0.399
-2.9	0.002	0.998	0.006	0.1	0.540	0.460	0.397
-2.8	0.003	0.997	0.008	0.2	0.579	0.421	0.391
-2.7	0.003	0.997	0.010	0.3	0.618	0.382	0.381
-2.6	0.005	0.995	0.014	0.4	0.655	0.345	0.368
-2.5	0.006	0.994	0.018	0.5	0.691	0.309	0.352
-2.4	0.008	0.992	0.022	0.6	0.726	0.274	0.333
-2.3	0.011	0.989	0.028	0.7	0.758	0.242	0.312
-2.2	0.014	0.986	0.035	0.8	0.788	0.212	0.290
-2.1	0.018	0.982	0.044	0.9	0.816	0.184	0.266
-2.0	0.023	0.977	0.054	1.0	0.841	0.159	0.242

تابع جدول (٨-١).

k	$F(k)$	$H(k)$	$f(k)$	k	$F(k)$	$H(k)$	$f(k)$
-1.9	0.029	0.971	0.066	1.1	0.864	0.136	0.218
-1.8	0.036	0.964	0.079	1.2	0.885	0.115	0.194
-1.7	0.045	0.955	0.094	1.3	0.903	0.097	0.171
-1.6	0.055	0.945	0.111	1.4	0.919	0.081	0.150
-1.5	0.067	0.933	0.130	1.5	0.933	0.067	0.130
-1.4	0.081	0.919	0.150	1.6	0.945	0.055	0.111
-1.3	0.097	0.903	0.171	1.7	0.955	0.045	0.094
-1.2	0.115	0.885	0.194	1.8	0.964	0.036	0.079
-1.1	0.136	0.864	0.218	1.9	0.971	0.029	0.066
-1.0	0.159	0.841	0.242	2.0	0.977	0.023	0.054
-0.9	0.184	0.816	0.266	2.1	0.982	0.018	0.044
-0.8	0.212	0.788	0.290	2.2	0.986	0.014	0.035
-0.7	0.242	0.758	0.312	2.3	0.989	0.011	0.028
-0.6	0.274	0.726	0.333	2.4	0.992	0.008	0.022
-0.5	0.309	0.691	0.352	2.5	0.994	0.006	0.018
-0.4	0.345	0.655	0.368	2.6	0.995	0.005	0.014
-0.3	0.382	0.618	0.381	2.7	0.997	0.003	0.010
-0.2	0.421	0.579	0.391	2.8	0.997	0.003	0.008
-0.1	0.460	0.540	0.397	2.9	0.998	0.002	0.006
				3.0	0.999	0.001	0.004



شكل (٨-٢). $F(k)$ ، و $f(k)$ و k ، و z **للتوزيع الطبيعي القياسي.**

المثال (٨-١). يوجد لدى الباحث بيانات عن التوزيع الطبيعي بحيث أن المتوسط 50 والانحراف المعياري 4 ويسعى للحصول على متغير أقل من 60 أو مساوياً له وبالتالي فإن

$$X \sim N(50, 4^2)$$
$$z = (60-50)\big/4 = 2.5$$

وينتج عن البحث في الجدول (٨-١) عندما يكون $z = 2.5$ ما يلي:

$$F(2.5) = 0.994$$

وبالتالي

$$P(X \leq 60) = F(2.5) = 0.994$$

المثال (٨-٢). علـى افتراض أن المتغير الطبيعي هو ذات المتغير المبيّن في المثال (٨-١) ويسعى الباحث لإيـجاد احتمال المتغير الواقع بين 42 و 58. ولتحقيق ذلك يتم اتباع الخطوات الخمس التالية:

١. $z_1 = (58 - 50)/4 = 2.0$
٢. $z_2 = (42 - 50)/4 = -2.0$
٣. $F(2.0) = 0.977$
٤. $F(-2.0) = 0.023$
٥. $P(42 \leq x \leq 58) = F(2.0) - F(-2.0) = 0.954$

٨-٦ بيانات العيّنة

عند توفر بيانات العيّنة $(x_1,\ldots,x_n)$، يتم حساب متوسط العيّنة $\overline{x}$، وتباينها s^2، وانحرافها المعياري s، على النحو التالي:

$$\overline{x} = \sum_{i=1}^{n} x_i \Big/ n$$
$$s^2 = \sum_{i=1}^{n} (x_i - \overline{x}) \Big/ (n-1)$$
$$s = \sqrt{s^2}$$

المثال (٨-٣). يوجد لدى المحلل المشاهدات المختارة عشوائياً $n = 10$:

136,99,106,89,111,105,87,94,109,123

يتم حساب متوسط العينة، وتباينها، وانحرافها المعياري على النحو المبين أدناه:

$$\begin{aligned}\overline{x} &= 1059/10 = 105.9\\ s^2 &= 2066.9/9 = 117.66\\ s &= \sqrt{117.66} = 10.85\end{aligned}$$

٨–٧ تقدير المعالم عند توفر بيانات العيّنة

عند توفر بيانات العينة يمثل متوسط العينة $\overline{x}$، وانحرافها المعياري s تقديراً لمتوسط المجتمع μ والانحراف المعياري σ على التوالي.

المثال (٨–٤). باستخدام البيانات المبيّنة في المثال (٨–٣)، يصبح تقدير المتوسط والانحراف المعياري على النحو التالي:

$$\begin{aligned}\hat{\mu} &= \overline{x} = 105.9\\ \hat{\sigma} &= s = 10.85\end{aligned}$$

على افتراض أنه قد تم توزيع البيانات طبيعياً بحيث أن

$$X \sim N(105.9, 10.85^2)$$

٨–٨ تقدير المعالم عند عدم توفر بيانات

في حال لم يتوفر لدى المحلل بيانات العينة لتقدير قيم المعالم، يتم استدعاء أحد الخبراء لتقديم بعض القياسات التقريبية المتعلقة بالتوزيع. وفيما يلي نوضح طريقتين لذلك:

تتطلب الطريقة الأولى القياسات التقريبية المدرجة أدناه:

$x\alpha_1 =$ النقطة المئوية α_1 للمتغير X

$x\alpha_2 =$ النقطة المئوية α_2 للمتغير X

من خلال البيانات المبيّنة أعلاه، يتم إدراج المعادلات التالية:

$$\begin{aligned}x\alpha_1 &= \mu + z\alpha_1\sigma\\ x\alpha_2 &= \mu + z\alpha_2\sigma\end{aligned}$$

حيث إن

$z\alpha_1 =$ النقطة المئوية α_1 على $Z \sim N(0,1)$

$z\alpha_2 =$ النقطة المئوية α_2 على $Z \sim N(0,1)$

باستخدام بعض العمليات الجبرية، يكون تقدير μ و σ على النحو المبين أدناه:

$$\hat{\sigma} = (x\alpha_2 - x\alpha_1)\big/(z\alpha_2 - z\alpha_1)$$
$$\hat{\mu} = x\alpha_1 - z\alpha_1\hat{\sigma}$$

المثال (٨-٥). يقوم الباحث بتصميم نموذج محاكاة ويحتاج إلى التوزيع الطبيعي، إلا أنه لا توجد لديه بيانات العينة. يتم إعطاء التقريبات التالية فيما يتعلق بالمتغير العشوائي:

$$(x\alpha_1, \alpha_1) = (50, 0.05)$$
$$(x\alpha_2, \alpha_2) = (100, 0.95)$$

ومن خلال البحث في الجدول (٨-١) لإيجاد قيم z التي تناظر $\alpha_1 = F(z_{0.05}) = 0.05$ و $\alpha_2 = F(z_{0.95}) = 0.95$ نجدها على النحو التالي:

$$z_{0.05} = -1.645$$
$$z_{0.95} = 1.645$$

وأخيراً، يكون تقدير الانحراف المعياري الطبيعي والمتوسط على النحو المبين أدناه:

$$\hat{\sigma} = \big[100 - 50\big]\big/\big[1.645 - (-1.645)\big] = 15.20$$
$$\hat{\mu} = 50 - (-1.645) \times 15.20 = 75.00$$

تتطلب الطريقة الثانية التقريب المتعلق بقياسات التوزيع التالية:

$L =$ الحد الأدنى

$H =$ الحد الأعلى

باعتبار أن المدى بين L و H في التوزيع الطبيعي هو ستة انحرافات معيارية، والمتوسط بين L و H، يكون تقدير المعلمتين على النحو التالي:

$$\hat{\mu} = (L + H)\big/2$$
$$\hat{\sigma} = (H - L)\big/6$$

المثال (٨-٦). يريد الباحث تطبيق التوزيع الطبيعي في إحدى الدراسات إلا أنه لا توجد لديه بيانات العينة. ولكن تتوفر لديه أفضل التقديرات المتعلقة بالحد الأدنى والأعلى على النحو التالي:

$$L = 10$$
$$H = 90$$

وبالتالي يصبح تقدير المعلمتين على النحو المبين أدناه:

$$\hat{\mu} = (10 + 90)\big/2 = 50$$
$$\hat{\sigma} = (90 - 10)\big/6 = 13.3$$

٨-٩ ملخص

يتّخذ التوزيع الطبيعي شكل الجرس وله المتغير العشوائي X والمعلمتان μ و σ . في حين أن للتوزيع الطبيعي القياسي المتغير العشوائي Z ويكون متوسطه صفراً وانحرافه المعياري واحداً. تتضمن معظم كتب الإحصاء جداول عن التوزيع الطبيعي القياسي. يعرض الفصل طريقة سهلة للتحويل من X إلى Z ومن Z إلى X . ونظراً لعدم وجود صيغة محددة للاحتمال التراكمي $F(z)$ ، قدم الفصل صيغة تقريبية. كما أنه توجد صيغة تقريبية لإيجاد Z من أحد قيم $F(z)$. عند توفر بيانات العينة، يتم استخدام متوسط العينة وانحرافها المعياري لتقدير متوسط التوزيع الطبيعي وانحرافه المعياري. وعند عدم توفر بيانات العينة، يتم تقديم القياسات التقريبية للتوزيع والتي تتيح تقدير متوسط التوزيع الطبيعي وانحرافه المعياري.

الفصل التاسع

توزيع اللوغاريتم الطبيعي
LOGNORMAL DISTRIBUTION

٩-١ مقدمة

قدم عالمان رياضيان بريطانيان، وهما فرانسيس غالتون ودونالد مكاليستر (Francis Galton and Donald McAlister)، توزيع اللوغاريتم الطبيعي عام 1879. ويُشار أحياناً إلى توزيع اللوغاريتم الطبيعي بتوزيع غالتون. كما يبدأ متغير توزيع اللوغاريتم الطبيعي عند الصفر، وتبلغ دالة كثافته ذروتها بعد فترةٍ وجيزة، ثم ينحدر إلى قيم x العليا و نقول عن توزيع المتغير بأنه يتبع توزيع اللوغاريتم الطبيعي عندما يتشكّل متغير آخر Y من خلال لوغاريتم X، بحيث يتوزّع توزيعاً طبيعياً. يناقش هذا الفصل دالة الكثافة الاحتمالية للمتغير X، في حين أنه لا يتطرق إلى دالة التوزيع التراكمية $F(x)$ باعتبار أنه لا توجد صيغة محددة لها. كما يتم تقديم طريقة لحساب الاحتمال التراكمي لأي متغير عشوائي X. عند توفر بيانات العينة، يتم استخدام قياسات العينة لتقدير معالم توزيع اللوغاريتم الطبيعي. وفي حال عدم توفر أي بيانات وهناك حاجة إلى تقدير متغير اللوغاريتم الطبيعي، فقد تم توضيح طريقتين لحساب التقدير. إنَّ توزيع اللوغاريتم الطبيعي ليس معروفاً مثل التوزيع الطبيعي، إلا أنه يُطبّق بسهولة في الدراسات البحثية بجميع أنواعها. وله تطبيقات في العديد من التخصصات، مثل الطقس، والهندسة ، والاقتصاد.

٩-٢ أساسيات

إن المتغير X يتبع توزيع لوغاريتم طبيعي عندما يتم توزيع اللوغاريتم الطبيعي $\ln$ بالنسبة للمتغير X توزيعاً طبيعياً. إن العلاقة بين x و y تكون على النحو التالي:

$$Ln(x) = y$$
$$e^{y} = x$$

بحيث إن معلمتي y هما المتوسط μ_y والانحراف المعياري σ_y ، ومعلمتي x هما μ_x و σ_x . إن رموز توزيع اللوغاريتم الطبيعي بالنسبة للمتغير X ، ورموز التوزيع الطبيعي المقابل Y تكون على النحو التالي مع ملاحظة أن المعلمتين اللتين تصفان توزيع المتغير X هما متوسط Y وانحرافه المعياري.

$$X \sim LN(\mu_y, \sigma_y^2)$$
$$Y \sim N(\mu_y, \sigma_y^2)$$

كما تم إدراج دالة الكثافة الاحتمالية $f(x)$ للمتغير X أدناه إلا أنه لم يتم ذكر دالة التوزيع التراكمية $F(x)$ وذلك لأنه لا توجد صيغة محددة للتوزيع التراكمي.

$$f(x) = \sqrt{\left(\frac{1}{2\pi x^2 \sigma_y^2}\right)}\; e^{-0.5\left[\frac{(y-\mu_y)}{\sigma_y}\right]^2}$$

إن العلاقة بين معلمتي x و y تكون على النحو التالي:

$$\mu_x = \exp\left[\mu_y + \sigma_y^2/2\right]$$
$$\sigma_x^2 = \exp\left[2\mu_y + \sigma_y^2\right]\left[\exp\left(\sigma_y^2\right) - 1\right]$$
$$\mu_y = Ln\left[\mu_x^2 / \sqrt{\mu_x^2 + \sigma_x^2}\right]$$
$$\sigma_y^2 = Ln\left[1 + \sigma_x^2/\mu_x^2\right]$$

٩-٣ منوال اللوغاريتم الطبيعي

يتم الحصول على منوال متغير اللوغاريتم الطبيعي X المُشار إليه بالرمز $\tilde{\mu}_x$ ، كما هو مبين أدناه:

$$\tilde{\mu}_x = exp(\mu_y - \sigma_y^2)$$

٩-٤ وسيط اللوغاريتم الطبيعي

يتم الحصول على وسيط الوغاريتم الطبيعي X على النحو التالي:

$$\mu_{0.5} = \exp(\mu_y)$$

المثال (٩-١). في تجربةٍ مخبرية، تم الكشف عن أحد المتغيرات X على أنه لوغاريتم طبيعي حيث إن $LN(2.5,1.1^2)$ بحيث يتم حساب متوسط وتباين X على النحو التالي:

$$\begin{aligned}
\mu_x &= \exp(2.5 + 1.1^2/2) = 22.31 \\
\sigma_x^2 &= \exp(2 \times 2.5 + 1.1^2)\left[\exp(1.1^2) - 1\right] = 1171.3
\end{aligned}$$

ومن ثم يتم الحصول على المنوال والوسيط على النحو التالي:

$$\begin{aligned}
\tilde{\mu}_x &= \exp(2.5 - 1.1^2) = 3.63 \\
\mu_{0.5} &= \exp(2.5) = 12.18
\end{aligned}$$

على افتراض أن الباحث يسعى للحصول على احتمال أن يكون X أقل من 30 أو مساوياً له. لتحقيق ذلك يتم إجراء الحسابات التالية:

$$\begin{aligned}
P(X \le 30) &= P\left[y \le \ln(30)\right] \\
&= P(y \le 3.40) \\
&= P\left(z\left[\le (3.4 - 2.5)/1.1\right]\right) \\
&= P(z \le 0.82) \\
&= F_z(0.82) = 0.794 \\
&= 0.794
\end{aligned}$$

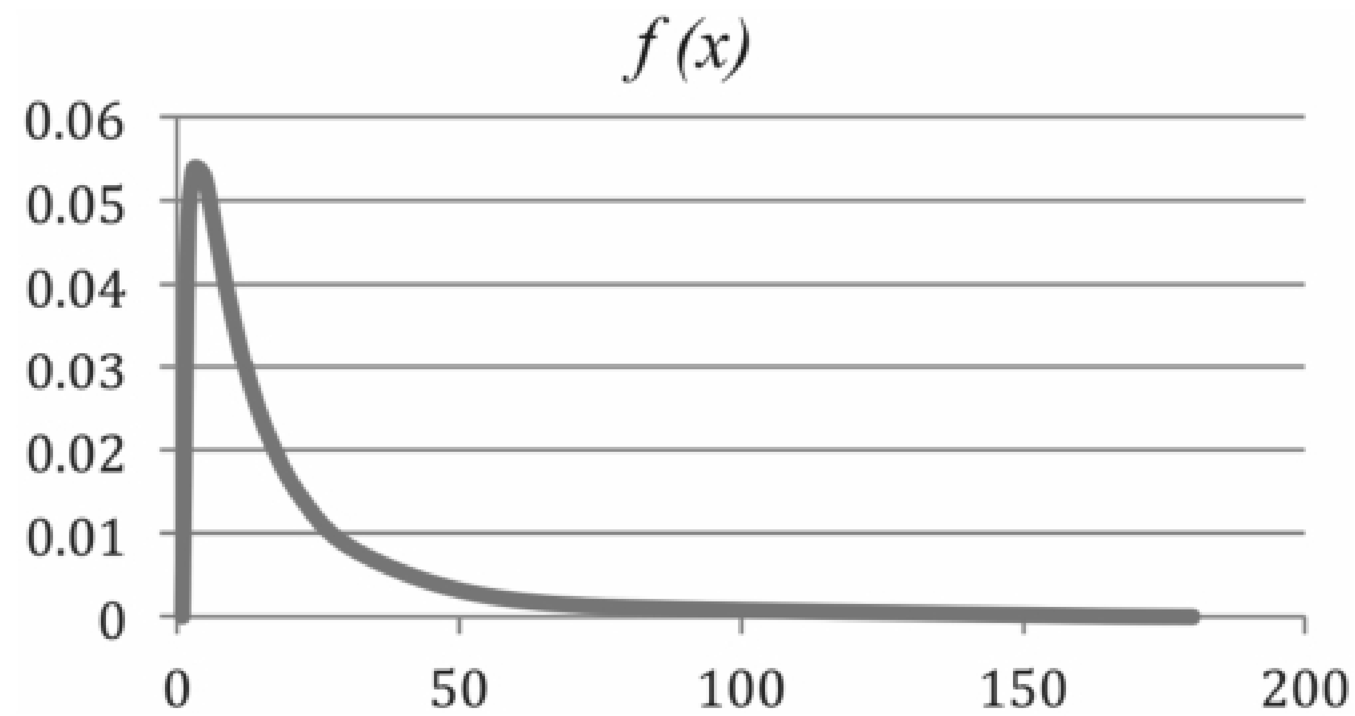
حيث إن $F_z(z)$ هو الاحتمال التراكمي للمتغير Z والذي تم إدراجه في الجدول (٨-١)

بافتراض أيضاً أن الباحث يريد إيجاد النقطة 0.90% للمتغير X. يتم التوصل إلى ذلك من خلال إيجاد النقاط 0.90% للمتغير $Z \sim N(0,1)$ ؛ $X \sim N(2.5,1.1^2)$ ؛ و $X \sim LN(2.5,1.1^2)$ ، كما هو مبين أدناه:

$$z_{0.90} = 1.282$$
$$y_{0.90} = 2.5 + 1.282 \times 1.1 = 3.91$$
$$x_{0.90} = \exp(3.91) = 49.9$$

ويلاحظ أن $F(49.9) = P(x \leq 49.9) = 0.90$

كما تم تمثيل الخط البياني للتوزيع في الشكل (٩-١).



شكل (٩-١). توزيع اللوغاريتم الطبيعي عندما $X \sim LN(2.5, 1.1^2)$.

٩-٥ بيانات العيّنة

عند توفر بيانات العيّنة مثل: $(x_1, \ldots, x_n)$، بحيث يسعى المحلل لتطبيق توزيع اللوغاريتم الطبيعي، تكون الخطوة الأولى في تحويل البيانات إلى $(y_1, \ldots, y_n)$ المناظرة حيث إن $y_i = \mathrm{Ln}(x_i)$ و $i = 1{:}n$ وLn هو الوغاريتم الطبيعي. إذا ظهرت البيانات التي تم تحويلها على أنها موزّعة طبيعياً، فإن توزيع بيانات X الأصلية موزّعة توزيع اللوغاريتم الطبيعي. ومن ثم يتم حساب متوسط البيانات المحوّلة وانحرافها المعياري على النحو التالي:

$\bar{y}$ = متوسط العينة

s_y = الانحراف المعياري للعينة

المثال (٩-٢). يوجد لدى المحلل 13 ملاحظة عن بيانات العينة $(i = 1\text{-}13)$ كما هو مبيّن في الجدول (٩-١)، ويفترض أن يتم تمثيل البيانات على شكل توزيع اللوغاريتم الطبيعي. يبين الجدول قيم $y_i = Ln(x_i)$ بالنسبة لكل مشاهدة من المشاهدات الثلاث عشرة.

جدول (٩-١). بيانات العينة $(x_1, \ldots, x_{13})$ **والبيانات المقابلة** $i = 1 - 13 \quad y_i = \ln(x_i),$.

x	y
12	2.485
215	5.371
23	3.135
2	0.693
35	3.555
27	3.296
11	2.398
89	4.489
13	2.565
45	3.807
82	4.407
28	3.332
8	2.079

يتم حساب متوسط عينة الثلاثة عشر مدخلاً بالنسبة للمتغير Y، وكذلك تباين العينة وانحرافها المعياري على النحو المبين أدناه:

$$\bar{y} = \sum_{i=1}^{13} y_i/_{13} = 3.201$$

$$s_y{}^2 = \sum_{i=1}^{13} \frac{(y_i - \bar{y})^2}{12} = 1.454$$

$$s_y = 1.206$$

مع ملاحظة أن القيمة الدنيا $y(1)$، والقيمة العليا $y(13)$ لمدخلات y تكون على النحو التالي:

$$y(1) = 0.693$$

$$y(13) = 5.371$$

٩-٦ تقدير المعالم عند توفر بيانات العيّنة

عند توفر بيانات العينة، يتم حساب المتوسط $\overline{y}$ والانحراف المعياري s_y ليصبح تقدير معلمتي اللوغاريتم الطبيعي على النحو التالي:

$$\hat{\mu}_y = \overline{y}$$
$$\hat{\sigma}_y = s_y$$

يتم اجراء تحليل آخر لتحديد ما إذا كان توزيع مدخلات y لبيانات العينة n يشابه التوزيع الطبيعي. وبالإشارة إلى أنه إذا كانت مدخلات y طبيعية، فإن مدخلات x هي لوغاريتم طبيعي. تتمثل إحدى طرق التحقق من الإعتدالية في اختبار نسبة الانتشار المبيّنة في الفصل العاشر. ويتم حساب نسبة الانتشار على النحو المبين أدناه:

$$\theta = \left[\overline{y} - y(1)\right] \big/ \left[y(n) - \overline{y}\right]$$

حيث إن

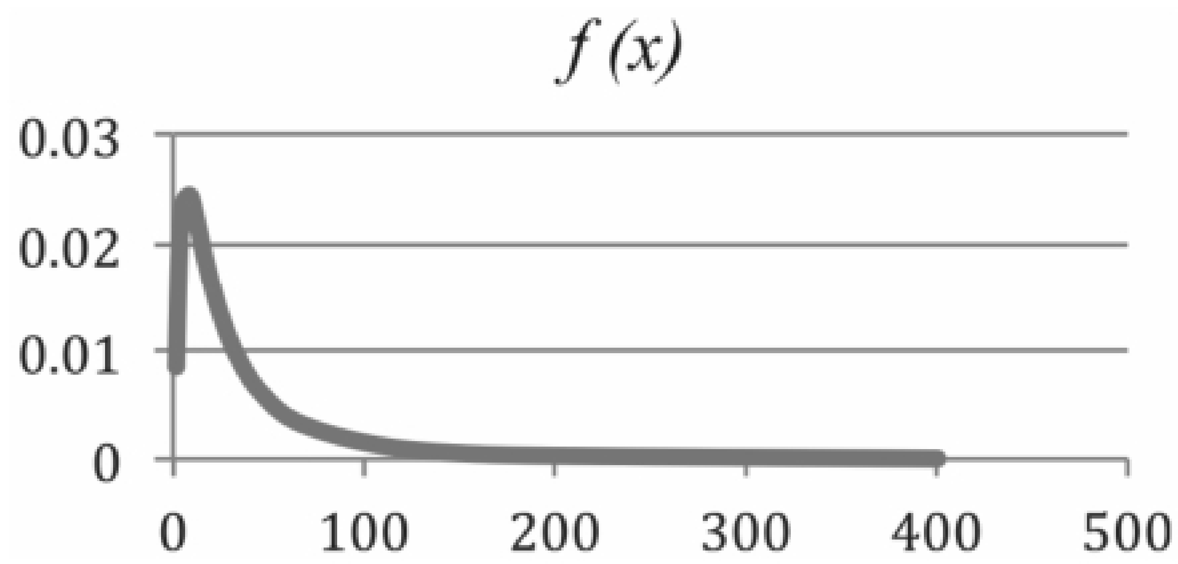
$$y(1) = \text{القيمة الدنيا} \left[y_1, \ldots, y_n\right]$$
$$y(n) = \text{القيمة العليا} \left[y_1, \ldots, y_n\right]$$

إذا كان θ قريباً من 1.00، يُعتبر توزيع مشاهدات y مشابهاً للتوزيع الطبيعي. إن المدى التقريبي لقبول التوزيع الطبيعي هو عندما تكون $(0.70 \leq \theta \leq 1.30)$

المثال (٩-٣). باستخدام بيانات عينة المثال (٩-٢)، يريد الباحث اختبار ما إذا كانت بيانات العينة x مماثلة بما يكفي لتوزيع اللوغاريتم الطبيعي. عندما يكون الأمر كذلك، يكون تقدير المعلمتين مناسباً للتطبيق في تحليل تطبيق اللوغاريتم الطبيعي. وللبدء في ذلك، يقوم الباحث بحساب نسبة انتشار θ كما هو مبين أدناه:

$$\begin{aligned} \theta &= (3.201 - 0.693) \big/ (5.371 - 3.201) \\ &= 1.15 \end{aligned}$$

باعتبار أن نسبة الانتشار قريبة من 1.00، يتم توزيع مشاهدات y بطريقة مماثلة للتوزيع الطبيعي. وبالتالي يُفترض أن تكون مشاهدات y على شكل التوزيع الطبيعي، وأن تكون مشاهدات x على شكل اللوغاريتم الطبيعي ويمثل الشكل (٩-٢) دالة توزيع اللوغاريتم الطبيعي.



شكل (٩-٢). لوغاريتم الطبيعي عندما $X \sim LN(3.201, 1.206^2)$.

وأخيراً يكون تقدير معلمتي اللوغاريتم الطبيعي على النحو التالي:

$$\hat{\mu}_y = \overline{y} = 3.201$$
$$\hat{\sigma}_y = s_y = 1.206$$

٩-٧ تقدير المعالم عند عدم توفر بيانات

عند عدم توفر بيانات العينة يسعى الباحث إلى تطبيق توزيع اللوغاريتم الطبيعي في إحدى الدراسات، فهناك طرق لتقدير المعالم عن طريق توفير عدة تقريبات فيما يتعلق بشكل توزيع اللوغاريتم الطبيعي. وقد تم توضيح طريقتين في هذا الفصل وهما:

١. تُطبق الطريقة الأولى عند إعطاء تقريبات النقطة 2٪ فيما يتعلق بشكل اللوغاريتم الطبيعي:

$(x\alpha_i, \alpha_i) =$ النقطة المئوية α_i للمتغير X

$(x\alpha_2, \alpha_2) =$ النقطة المئوية α_2 للمتغير X

استناداً إلى التوزيع الطبيعي القياسي، تكون النقط المئوية المقابلة للمتغير Z على النحو التالي:

$z\alpha_i =$ النقطة المئوية α_i للمتغير Z

$z\alpha_2 =$ النقطة المئوية α_2 للمتغير Z

مع ملاحظة العلاقتين الطبيعيتين أدناه اللتين تشكّلان معادلتين بمجهولين:

$$y\alpha_i = \mu_y + z\alpha_i\sigma_y$$
$$y\alpha_2 = \mu_y + z\alpha_2\sigma_y$$

وأخيراً، يتم التوصل إلى تقدير معلمتي اللوغاريتم الطبيعي كما هو مبين أدناه:

$$\hat{\sigma}_y = (y\alpha_2 - y\alpha_i)\big/(z\alpha_2 - z\alpha_i)$$
$$\hat{\mu}_y = (y\alpha_i - z\alpha_i\hat{\sigma}_y)$$

٢. تتطلب الطريقة الثانية لتقدير معلمتي اللوغاريتم الطبيعي عند عدم توفر بيانات العينة التقريبات التالية لتوزيع اللوغاريتم الطبيعي:

$\tilde{x} =$ القيمة المرجحة للمتغير X (المنوال)

$x_{0.5} =$ القيمة الوسطى للمتغير X (الوسيط)

مع ملاحظة علاقتي اللوغاريتم الطبيعي بالنسبة للمنوال والوسيط كما هو مبين أدناه:

$$\tilde{x} = e^{\mu_y - \sigma_y^2}$$
$$x_{0.5} = e^{\mu_y}$$

عند تطبيق بعض العمليات الجبرية، يتم تقدير معلمتي اللوغاريتم الطبيعي على النحو التالي:

$$\hat{\mu}_y = \ln(x_{0.5})$$
$$\hat{\sigma}_y = \left[\hat{\mu}_y - \ln(\tilde{x})\right]^{0.5}$$

المثال (٩-٤). يريد أحد المهندسين تطبيق اللوغاريتم الطبيعي في إحدى الدراسات إلا أنه ليس لديه بيانات العينة لتقدير قيم المعالم. وبإجراء بعض التحليل، أصبح لدى المهندس بعض التقريبات عن النقطتين المئويتين 2% للمتغير X :

$$x_{0.2} = 10$$
$$x_{0.8} = 70$$

يتم التوصل إلى النقط المئوية المقابلة للتوزيع الطبيعي على النحو المبين أدناه:

$$y_{0.2} = \ln(10) = 2.302$$
$$y_{0.8} = \ln(70) = 4.248$$

وإضافةً إلـى ذلك، تكون النقطتان المئويتان ذات الصلة للتوزيع الطبيعي القياسي علـى النحو التالي:

$$z_{0.2} = -0.841$$
$$z_{0.8} = 0.841$$

ويكون تقدير معلمتي اللوغاريتم الطبيعي على النحو التالي:

$$\hat{\sigma}_y = (4.248 - 2.302) \big/ \left[0.841 - (-0.841)\right] = 1.157$$
$$\hat{\mu}_y = 2.302 - (-0.841) \times 1.157 = 3.275$$

وأخيراً يكون للمتغير x توزيع اللوغاريتم الطبيعي كما هو مبيّن أدناه:

$$X \sim LN(3.275, 1.157^2)$$

المثال (٩-٥). في إحدى دراسات المحاكاة، يحتاج المحلل إلى تطبيق توزيع اللوغاريتم الطبيعي ولا يوجد لديه بيانات العينة إلا أنه يوجد لديه التقريبين التاليين للتوزيع:

$$\tilde{x} = 100$$
$$x_{0.5} = 200$$

باستخدام الطريقـة المبينة أعـلاه يكون تقدير معلمتي اللوغاريتم الطبيعـي علـى النحو التالي:

$$\hat{\mu}_y = \ln(200) = 5.298$$
$$\hat{\sigma}_y = \left[5.298 - \ln(100)\right]^{0.5} = 0.832$$

ويكون متغير اللوغاريتم الطبيعي الذي ينبغي استخدامـه فـي دراسة المحاكاة على النحو المبين أدناه:

$$X \sim LN(5.298, 0.832^2)$$

٩-٨ ملخص

يتم توزيع المتغير X بتوزيع اللوغاريتم الطبيعي عندما يتم توزيع المتغير الآخر Y والذي يمثل لوغاريتم X توزيعاً طبيعياً. ومن خلال بعض التحويلات البسيطة التي تربط X و Y، تم إدراج المعادلات التي تبيّن كيفية تحويل متوسط وتباين X إلى متوسط وتباين Y. إن المعلمتين اللتين تصف متغير اللوغاريتم الطبيعي X هما متوسط وتباين Y. كما تم إدراج دالة الكثافة الاحتمالية للمتغير X، إلا أنه لم يتم توضيح دالة التوزيع التراكمي باعتبار أنه لا توجد صيغة محددة للتوزيع التراكمي. في حال كان هناك ضرورة للاحتمال التراكمي لقيمة معيّنة للمتغير X، يتم تطبيق طريقة الحساب بسهولة. عند توفر بيانات العينة، يتم استخدام قياسات البيانات لتقدير معلمتي اللوغاريتم الطبيعي عند الحاجة. وعند عدم توفر بيانات العينة، يتم توضيح الطريقتين اللتين تعطيان تقدير معلمتي اللوغاريتم الطبيعي.

الفصل العاشر

التوزيع الطبيعي المقطوع الأيسر
LEFT TRUNCATED NORMAL DISTRIBUTION

١٠–١ مقدمة

في كتابٍ سابق لثوموبولوس [Thomopoulos (1980) p 318–324]، يوضح الكاتب كيفية استخدام التوزيع الطبيعي المقطوع الأيسر على تطبيقات في مراقبة المخزون. ويتّخذ التوزيع الطبيعي المقطوع الأيسر عدة أشكال تتراوح ما بين طبيعي إلى شبه أسي. كما أن للتوزيع معلمة واحدة k تمثل قيمة معيّنة للتوزيع الطبيعي القياسي Z، بحيث يتضمن التوزيع جميع قيم z التي تكون أكبر من k. يُشار إلى المتغير بأنه t حيث إن $t=(z-k)$ وباعتبار أن z يساوي k أو أكبر منها، فإن قيمة t هي صفر أو أكبر. إن القياسات الرئيسة للتوزيع هي المتوسط، والانحراف المعياري، ومعامل الاختلاف، وقياس جديد يُسمى نسبة الانتشار وهي دائماً أكبر من صفر. وهناك أيضا مقياس آخر هو $t\alpha =$ النقطة المئوية α للمتغير t التي تعطي قيمة t من خلال الاحتمال التراكمي α. كما أن هناك تركيزاً على إيجاد النقطتين $t_{0.01}$ و $t_{0.99}$ و لقد تم إدراج جميع هذه القياسات في جدول لقيم k التي تتراوح ما بين -3.0 إلى $+3.0$. وعندما يتوفر لدى المحلل بيانات العينة $(x_1,\ldots,x_n)$ بحيث يأخذ إحصاءات معينة من البيانات بما في ذلك تقدير نسبة الانتشار، يمكن للمحلل تحديد ما إذا كان التوزيع الطبيعي المقطوع الأيسر أفضل للبيانات من التوزيع الطبيعي. في حال تم اختيار أي من التوزيعين، يتم حساب الحد الأدنى، المشار إليه بـ γ، حيث إن $x \geq \gamma$. كما يبين الفصل كيفية حساب $x\alpha$ النقطة المئوية α للمتغير X بالنسبة لأي α.

١٠-٢ أساسيات

يتّخذ التوزيع الطبيعي المقطوع الأيسر عدة أشكال تتراوح ما بين طبيعي إلى شبه أسي، وله معلمة واحدة k بحيث تشكل التوزيع. يتم إجراء استعراض سريع للتوزيع الطبيعي القياسي ذي المتغير Z باعتبار أن جميع قيم Z تقع ما بين -3.0 إلى $+3.0$، حيث يتضمن التحليل في هذا الفصل هذا المدى فقط.

١٠-٣ التوزيع الطبيعي القياسي

ربما يكون التوزيع الطبيعي القياسي أكثر توزيع احتمالي يستخدمه الباحثون في جميع أنواع التخصصات حيث تم وصف هذا التوزيع بالتفصيل في الفصل الثامن. إن حدود المتغير العشوائي Z تقع بين $-\infty$ إلى $+\infty$، ويكون المتوسط صفراً والتباين واحداً، ويرمز له بالرمز:

$$Z \sim N(0,1)$$

وتكون القيمة المتوقعة للمتغير Z والتباين معرفتين على النحو التالي:

$$E(Z) = 0$$
$$V(Z) = 1$$

نورد فيما يلي دالة الكثافة الاحتمالية للمتغير Z، ودالة الاحتمال التراكمي، والاحتمال المكمّل حيث إنه لا توجد صيغة محددة للاحتمال التراكمي وقد تم وضع العديد من الطرق الكمّية لحساب $F(x)$. وقد تم توضيح إحدى هذه الطرق في الفصل الثامن بحيث يتم استخدامها في بعض حسابات هذا الفصل.

الكثافة الاحتمالية للمتغير Z = $f(z) = \left(1\big/\sqrt{2\pi}\right)e^{-z^2/2}$

الاحتمال التراكمي عندما $z = k$، $F(k) = P(z \le k)\int\limits_{-\infty}^{k} f(z)\,dz$

الاحتمال المكمّل عندما $z = k$، $H(k) = P(z > k) = 1 - F(k)$

مع ملاحظة أن:

$$\int\limits_{k}^{\infty} zf(z)\,dz = f(k)$$

١٠-٤ التوزيع الطبيعي المقطوع الأيسر

يستخدم المتغير الطبيعي القياسي Z لتعريف المتغير العشوائي للتوزيع الطبيعي المقطوع الأيسر t والمعلمة المشار إليها بالرمز k وهي قيمة معيّنة للمتغير Z . في هذه الحالة يكون z أكبر من k أو مساوياً لها، وبالتالي t أكبر من الصفر أو مساوياً له كما هو مبين أدناه:

$$
\begin{aligned}
&t = z - k \qquad\quad z \geq k \\
&t \geq 0
\end{aligned}
$$

ويرمز للتوزيع الطبيعي المقطوع الأيسر بالرمز التالي:

$$t \sim LTN(k)$$

كما تعرف دالة الكثافة الاحتمالية $g(t)$ ، ودالة التوزيع التراكمية $G(t)$ للمتغير t ، على النحو المبين أدناه:

$$
\begin{aligned}
g(t) &= f(z)\big/H(k) \\
G(t) &= \Big[F(z) - F(k)\Big] \,/\, H(k)
\end{aligned}
$$

فيما يلي نوضح طريقة حساب متوسط وتباين t ذي المعلمة k في هذا الفصل. هناك حاجة لمعرفة التوقعات الجزئية $E(z > k)$ ، و $E\Big[(z > k)^2\Big]$ ، ويتم التوصل إليهما كما هو مبين أدناه:

$$
\begin{aligned}
E(z > k) &= \int_k^{\infty} (z - k) f(z)\, dz = f(k) - kH(k) \\
E\Big[(z > k)^2\Big] &= \int_k^{\infty} (z - k)^2 f(z)\, dz = -kf(k) + H(k)(1 + k^2)
\end{aligned}
$$

نورد فيما يلي توضيحاً لكيفية التوصل إلى القيمة المتوقعة للمتغير t و t^2 باستخدام المعلمة k .

القيمة المتوقعة للمتغير t ذي المعلمة k $\; E(t)_k = E(z > k)\big/H(k) =$

قيمة t^2 المتوقعة بالمعلمة k $\; E(t^2)_k = E\Big[(z > k)^2\Big]\Big/H(k) =$

تباين t ذو المعلمة k $\; V(t)_k = E(t^2)_k - E(t)_k^2 =$

وأخيراً، يكون متوسط t وانحرافه المعياري على النحو التالي:

متوسط t ذو المعلمة k $\mu_t(k) = E(t)_k = k$

الانحراف المعياري للمتغير t ذي المعلمة k $\sigma_t(k) = \sqrt{V(t)_k} = k$

يتم التوصل إلى معامل اختلاف المتغير t ذي المعلمة k على النحو المبين أدناه:

$$\mathrm{cov}_t(k) = \sigma_t(k) \big/ \mu_t(k)$$

١٠-٥ الاحتمال التراكمي للمتغير العشوائي t

نبيّن فيما يلي كيفية إيجاد قيمة $t\alpha$ النقطة المئوية α للمتغير t . مع ملاحظة أن $G(t)$ هو الاحتمال التراكمي للمتغير $t\alpha$ بحيث يكون:

$$G(t\alpha) = P(t \le t\alpha) = \alpha$$

ومع ملاحظة العلاقة التالية بين α و k للتوزيع الطبيعي المقطوع الأيسر و $F(z)$ ، $F(k)$ ، $H(k)$ للتوزيع الطبيعي القياسي:

$$G(t) = \alpha = \left[F(z) - F(k)\right] / \, H(k)$$

باستخدام بعض العمليات الجبرية، تعطي النتائج ما يلي:

$$\begin{aligned}
(1-\alpha) &= 1 - \left[F(z) - F(k)\right] \big/ H(k) \\
&= \left[H(k) - F(z) + F(k)\right] \big/ H(k) \\
&= \left[1 - F(z)\right] \big/ H(k) \\
&= H(z) \big/ H(k)
\end{aligned}$$

وبالتالي يتم حساب الاحتمال المكمّل للمتغير z من التوزيع الطبيعي القياسي على النحو المبيّن أدناه:

$$H(z) = (1-\alpha)\, H(k)$$

عند توفر قيمة $H(z)$ تهدف الخطوة التالية إلى إيجاد قيمة z ، المُشار إليها بالرمز z' ، والتي تقابل $H(z)$ ومن ثم الحصول على $t\alpha = z' - k$. ويتضح ذلك في الخطوات الخمس أدناه:

١. من k، نوجد $H(k)$ باستخدام الجدول (٨-١)

٢. $H(z') = (1-\alpha)\, H(k)$

٣. $F(z') = \left[1 - H(z')\right]$

٤. نوجد z' باستخدام الجدول (٨-١)

٥. $t\alpha = z' - k$

عندما يكون $\alpha = 0.01$ و 0.99، يتم التوصل إلى النقطتين المئويتين التاليتين:

$0.01 = t_{0.01}$ النقطة المئوية

$0.99 = t_{0.99}$ النقطة المئوية

مع ملاحظة أن متوسط t وانحرافه المعياري هما على النحو التالي:

$\mu_{t(x)} =$ متوسط المتغير t ذي المعلمة k

$\sigma_{t(x)} =$ الانحراف المعياري للمتغير t ذي المعلمة k

إن نسبـة الانتشار المُشار إليها بالرمز θ هي وحيدة لكل معلمة k، ويتم حسابها كما هو مبين أدناه:

$$\theta = \left[\mu_{t(k)} - t_{0.01}\right] \Big/ \left[t_{0.99} - \mu_{t(k)}\right] = \text{نسبة الانتشار}$$

يبيّن الجدول (١٠-١) مجمـوعة القياسات ($\mu_{t(k)}$، و $\sigma_{t(k)}$ ، $\mathrm{cov}_{t(k)}$، و $t_{0.01}$ ، $t_{0.99}$، و θ) للمتغير t من التوزيع الطبيعي المقطوع الأيسر حيث $k = \left[-3.0, (0,1), 3.0\right]$

جدول (١٠-١). التوزيع الطبيعي المقطوع الأيسر من خلال معلمة الموقع k، والنقطتين المئويتين $t_{0.01}$ و $t_{0.99}$، والمتوسط $\mu_t(k)$، والانحراف المعياري $\sigma_t(k)$، ومعامل الاختلاف، $\mathrm{cov}_t(k)$ ونسبة الانتشار θ.

k	$t_{0.01}$	$t_{0.99}$	$\mu_{t(k)}$	$\sigma_{t(k)}$	$\mathrm{cov}(k)$	θ
-3.0	0.720	5.327	3.004	0.993	0.33	0.98
-2.9	0.637	5.227	2.906	0.991	0.34	0.98
-2.8	0.559	5.127	2.808	0.989	0.35	0.97
-2.7	0.486	5.028	2.710	0.986	0.36	0.96
-2.6	0.419	4.928	2.614	0.982	0.38	0.95
-2.5	0.359	4.829	2.518	0.978	0.39	0.93
-2.4	0.305	4.729	2.423	0.972	0.40	0.92

تابع جدول (١٠-١).

k	$t_{0.01}$	$t_{0.99}$	$\mu_{t(k)}$	$\sigma_{t(k)}$	$\text{cov}(k)$	θ
-2.3	0.258	4.630	2.329	0.966	0.41	0.90
-2.2	0.218	4.532	2.236	0.959	0.43	0.88
-2.1	0.183	4.433	2.145	0.951	0.44	0.86
-2.0	0.155	4.335	2.055	0.942	0.46	0.83
-1.9	0.130	4.237	1.968	0.931	0.47	0.81
-1.8	0.110	4.140	1.882	0.920	0.49	0.78
-1.7	0.093	4.043	1.798	0.907	0.50	0.76
-1.6	0.080	3.947	1.717	0.894	0.52	0.73
-1.5	0.068	3.852	1.639	0.879	0.54	0.71
-1.4	0.059	3.758	1.563	0.863	0.55	0.69
-1.3	0.051	3.664	1.490	0.847	0.57	0.66
-1.2	0.044	3.572	1.419	0.830	0.58	0.64
-1.1	0.039	3.481	1.352	0.812	0.60	0.62
-1.0	0.034	3.390	1.288	0.794	0.62	0.60
-0.9	0.030	3.302	1.226	0.775	0.63	0.58
-0.8	0.027	3.214	1.168	0.756	0.65	0.56
-0.7	0.024	3.128	1.112	0.736	0.66	0.54
-0.6	0.022	3.044	1.059	0.717	0.68	0.52
-0.5	0.020	2.962	1.009	0.697	0.69	0.51
-0.4	0.018	2.881	0.962	0.678	0.70	0.49
-0.3	0.017	2.802	0.917	0.659	0.72	0.48
-0.2	0.015	2.724	0.875	0.640	0.73	0.47
-0.1	0.014	2.649	0.835	0.621	0.74	0.45
0.0	0.012	2.576	0.798	0.603	0.76	0.44
0.1	0.011	2.504	0.763	0.585	0.77	0.43
0.2	0.010	2.435	0.729	0.568	0.78	0.42
0.3	0.010	2.367	0.698	0.551	0.79	0.41
0.4	0.009	2.301	0.669	0.534	0.80	0.40
0.5	0.008	2.238	0.641	0.518	0.81	0.40
0.6	0.008	2.176	0.615	0.503	0.82	0.39
0.7	0.007	2.117	0.590	0.488	0.83	0.38
0.8	0.007	2.059	0.567	0.473	0.83	0.38

تابع جدول (١٠-١).

k	$t_{0.01}$	$t_{0.99}$	$\mu_{t(k)}$	$\sigma_{t(k)}$	$\text{cov}(k)$	θ
0.9	0.007	2.003	0.546	0.460	0.84	0.37
1.0	0.007	1.949	0.525	0.446	0.85	0.36
1.1	0.006	1.897	0.506	0.433	0.86	0.36
1.2	0.006	1.846	0.488	0.421	0.86	0.35
1.3	0.006	1.797	0.470	0.409	0.87	0.35
1.4	0.006	1.750	0.454	0.398	0.88	0.35
1.5	0.005	1.704	0.439	0.387	0.88	0.34
1.6	0.005	1.660	0.424	0.376	0.89	0.34
1.7	0.005	1.617	0.410	0.366	0.89	0.34
1.8	0.005	1.575	0.397	0.356	0.90	0.33
1.9	0.005	1.534	0.385	0.347	0.90	0.33
2.0	0.004	1.495	0.373	0.338	0.91	0.33
2.1	0.004	1.456	0.362	0.330	0.91	0.33
2.2	0.004	1.417	0.351	0.321	0.91	0.33
2.3	0.004	1.379	0.341	0.313	0.92	0.33
2.4	0.004	1.340	0.332	0.306	0.92	0.33
2.5	0.003	1.301	0.323	0.298	0.92	0.33
2.6	0.003	1.260	0.314	0.291	0.93	0.33
2.7	0.003	1.218	0.306	0.284	0.93	0.33
2.8	0.002	1.173	0.298	0.277	0.93	0.34
2.9	0.002	1.124	0.291	0.270	0.93	0.35
3.0	0.001	1.070	0.283	0.264	0.93	0.36

المثال (١٠-١). تبيّن الخطوات الخمس التالية الحسابات اللازمة لإيجاد $t_{0.01}$ عندما تكون $k = 1.0$.

١. $H(1.0) = 0.159$ من الجدول (٨-١)

٢. $H(z') = (1 - 0.01) \times 0.159 = 0.157$

٣. $F(z') = 1 - 0.157 = 0.843$

٤. $z' \approx 1.008$ عبر الجدول (٨-١)

٥. $t_{0.01} \approx 1.008 - 1.000 = 0.008$

بينما نلاحظ أن الجدول (١٠-١) يبيّن أن $t_{0.01} = 0.007$ عندما تكون $k = 1.0$. يعود سبب الفرق بين الجدول والحسابات المبينة هنا إلى التقريب.

المثال (١٠-٢). تبين الخطوات الخمس التالية الحسابات المتعلقة بـ $t_{0.99}$ عندما تكون $k = 1.0$.

١. $H(1.0) = 0.159$ من الجدول (٨-١)

٢. $H(z') = (1 - 0.99) \times 0.159 = 1.0016$

٣. $F(z') = 1 - 0.0016 = 0.9984$

٤. $z' \approx 2.94$ عبر الجدول (٨-١)

٥. $t_{0.01} \approx 2.94 - 1.00 = 1.94$

نلاحظ أن الجدول (١٠-١) يبين أن $t_{0.99} = 1.949$ عندما تكون $k = 1.0$. ويعود سبب الفرق بين الجدول والحسابات المبينة هنا إلى التقريب أيضاً.

المثال (١٠-٣). أوجد الحسابات اللازمة للحصول على $\mu_t(k)$، والانحراف المعياري $\sigma_t(k)$، ومعامل الاختلاف، $\text{cov}(k)$، ونسبة الانتشار θ بالنسبة للمتغير t عندما تكون $k = 1.0$.

يتم حساب المتوسط، والانحراف المعياري، ومعامل الاختلاف على النحو المبين أدناه:

$$
\begin{aligned}
\mu_t(1.0) &= \left[f(1.0) - 1.0 \times H(1.0)\right] \big/ H(1.0) \\
&= \left[0.242 - 1.0 \times 0.159\right] \big/ 0.159 \\
&= 0.522 \\
E(t^2)_{1.0} &= \left[-1.0 f(1.0) + H(1.0)(1 + 1.0^2)\right] \big/ H(1.0) \\
&= \left[-1 \times 0.242 + 0.159 \times (1 + 1)\right] \big/ 0.159 \\
&= 0.478 \\
V(t)_{1.0} &= 0.478 - 0.522^2 \\
&= 0.453 \\
\sigma_t(k) &= \sqrt{0.205} \\
&= 0.453 \\
\text{cov}(1.0) &= 0.453 \big/ 0.522 \\
&= 0.868
\end{aligned}
$$

نلاحظ الفرق في تقريب قيم الجدول المدرجة أدناه عندما تكون $k = 1.0$:

$$\begin{aligned}\mu_t(k) &= 0.525\\ \sigma_t(k) &= 0.446\\ \mathrm{cov}(k) &= 0.85\end{aligned}$$

كما يتم الحصول على نسبة الانتشار أدناه باستخدام قيم الجدول بدلاً من القيم المحسوبة.

$$\begin{aligned}\theta &= (0.525 - 0.007)\Big/(1.949 - 0.525)\\ &= 0.363\end{aligned}$$

١٠-٦ بيانات العيّنة

عند توفر بيانات العينة $(x_1, \ldots, x_n)$، يمكن للمحلل التوصل إلى القياسات الإحصائية التالية فيما يتعلق بالمتغير X :

$\overline{x} =$ المتوسط

$s =$ الانحراف المعياري

$$x(1) = \text{القيمة الدنيا } (x_1, \ldots, x_n)$$

$$x(n) = \text{القيمة العليا } (x_1, \ldots, x_n)$$

١٠-٧ تقدير المعالم عند توفر بيانات العيّنة

بعد جمع القياسات الإحصائية عن بيانات العينة، يتم حساب تقدير نسبة الانتشار المُشار إليها بالرمز $\hat{\theta}$ على النحو التالي:

$$\hat{\theta} = \left[\overline{x} - x(1)\right]\Big/\left[x(n) - \overline{x}\right]$$

وبعد تقدير نسبة الانتشار، يبحث المحلل في الجدول (١٠-١) للتوصل إلى القياسات التالية في التوزيع الطبيعي المقطوع الأيسر:

$k =$ معلمة التوزيع الطبيعي المقطوع الأيسر

$\mu_t(k) =$ متوسط t عند k

$\sigma_t(k) =$ الانحراف المعياري للمتغير t عند k

$0.01 = t_{0.01} =$ النقطة المئوية 0.01 للمتغير t عند k

كقاعدة عامة، إذا كان $\hat{\theta}$ في المدى الذي يتراوح ما بين 0.70 و 1.00 فلا تختلف بيانات العينة كثيراً عن بيانات التوزيع الطبيعي ويُطبّق التوزيع الطبيعي. أما إذا $\hat{\theta} < 0.70$ ، يعتبر التوزيع الطبيعي المقطوع الأيسر ذو المعلمة k التوزيع الأفضل لبيانات العينة.

عند اختيار التوزيع الطبيعي المقطوع الأيسر وعند تحديد k، يصبح شكل التوزيع أوضح وهو ما يتيح للمحلل تقدير الحد الأدنى للمتغير x ، وكذلك تقدير مختلف قيم النقطة المئوية α له. وسوف نوضح ذلك فيما يلي:

أولاً: يتم التوصل إلى تقدير الحد الأدنى للمتغير x باستخدام المعادلة التالية:

$$\hat{\gamma} = \overline{x} + s\left[\left(t_{0.01} - \mu_t(k)\right)\big/\sigma_t(k)\right]$$

ثانياً: يتم التوصل إلى تقدير النقطة المئوية α للمتغير X . وللحصول عليه، هناك حاجة إلى قيمة النقطة المئوية α للمتغير t عند k. ويتم الحصول على هذه القيمة من خلال تطبيق الخطوات الخمس التي تم ذكرها سابقاً في هذا الفصل. و بعد أن أصبح $t\alpha$ معروفاً، يتم حساب تقدير $x\alpha$ باستخدام المعادلة التالية:

$$x\alpha = \overline{x} + s\left[\left(t\alpha - \mu_t(k)\right)\big/\sigma_t(k)\right]$$

المثال (١٠-٤). يوجد لدى محلل n عينة عشوائية من تجربة ما وحصل على القياسات الإحصائية المبينة أدناه.

$$\overline{x} = 100$$
$$s = 25$$
$$x(1) = 70$$
$$x(n) = 150$$

يعتقد المحلل أن البيانات ليست موزّعة توزيعاً طبيعياً ويسعى إلى تطبيق التوزيع الطبيعي المقطوع الأيسر. يتم حساب تقدير نسبة الانتشار على النحو المبين أدناه:

$$\hat{\theta} = (100 - 70)\big/(150 - 100) = 0.60$$

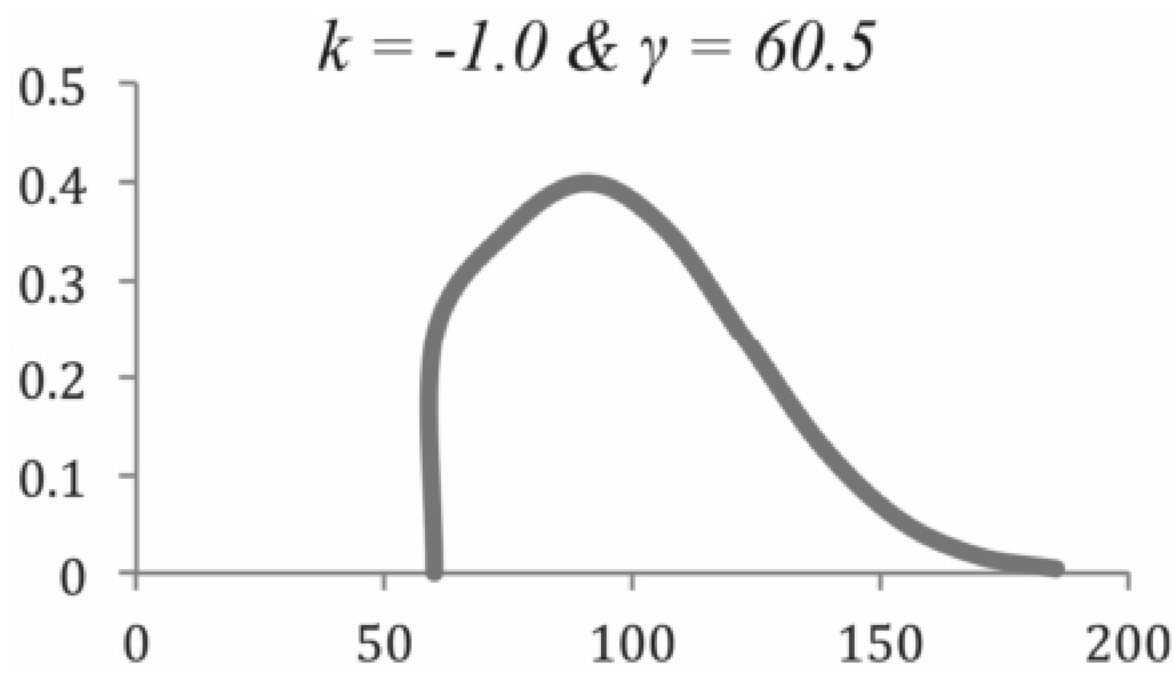
ومن ثم يتم تطبيق نسبة الانتشار المقدّرة في الجدول (١٠-١) لإيجاد القيمة الجدولية الأقرب للنسبة θ. عندما تكون $\theta = 0.60$، تعطي القيمة الجدولية $k = -1.00$، والقياسات التالية للتوزيع الطبيعي المقطوع الأيسر:

$$t_{0.01} = 0.034$$
$$\mu_t(k) = 1.288$$
$$\sigma_t(k) = 0.794$$

من خلال هذه القياسات، تم تقدير الحد الأدنى للمتغير x كما هو مبين أدناه:

$$\begin{aligned}\hat{\gamma} &= 100 + 25(0.034 - 1.288)\big/0.794 \\ &= 60.5\end{aligned}$$

يمثل الشكل (١٠-١) الرسم البياني للتوزيع الطبيعي المقطوع الأيسر، حيث إن $k = -1.00$ و $\gamma = 60.5$.



شكل (١٠-١). التوزيع الطبيعي المقطوع الأيسر حيث إن $k = -1.00$ و $\gamma = 60.0$.

المثال (١٠-٥). باستخدام بيانات العينة المبينة في المثال (١٠-٤)، وبافتراض أن المحلل يسعى للحصول على قيمة x من خلال الاحتمال التراكمي 0.95، والتي تمثل النقطة المئوية $x_{0.95}$ للمتغير X ويتضح ذلك في الخطوات الست التالية:

١. نوجد أولاً $H(-1.00)=0.841$ من الجدول (٨-١)

٢. $H(z')=(1-0.95)0.841=0.042$

٣. $F(z')=(1-0.042)=0.958$

٤. $z'=1.73$ من الجدول (٨-١)

٥. $t_{0.95}=1.73-(-1.00)=2.73$

٦. $x_{0.95}=100+25\left[(2.73-1.288)/0.794\right]=145.4$

و بالتالي

$$P(X\leq 145.4)\approx 0.95$$

٨-١٠ التوزيع الطبيعي المقطوع الأيسر في ضبط المخزون

يعدُّ ضبط المخزون، وكذلك الوقت المناسب للشراء ومقداره دالة مهمة في جميع مراكز التوزيع، وتجار التجزئة، والوكلاء، والمتاجر. يحتوي مركز التوزيع النموذجي على 100000 رقم قطعة ويوجد لدى الوكيل 20000 ويتم وضع التوقعات F كل شهر بناءً على طلب كل قطعة في كل موقع. كما يتم حساب الانحراف المعياري لخطأ التوقع s، ويتم قياس معامل الاختلاف (cov) على أنه $\text{cov}=s/F$. إن معامل الاختلاف هو قياس خطأ التوقع شهرياً بالنسبة لكل قطعة ويتم تحديثه كل شهر. إن الهدف من إدارة المخزون هو توفير الحد الأدنى من المخزون اللازم لتحقيق مستوى الخدمة المطلوبة (SL) حيث إن $SL=$ الطلب المنفّذ/ إجمالي الطلب، و مستوى الخدمة المثالي المطلوب هو 0.95. كما أن هناك حاجة للتوقع ولأخطاء التوقع لتحديد المقدار المحدد للمخزون الذي ينبغي توفيره. وفي معظم نُظم المخزون، يُفترض التوزيع الطبيعي على أنه الطلبات الشهرية لكل طرف وهذا ليس صحيح، حيث إن التوزيع الأكثر شيوعاً للطلبات الشهرية هو التوزيع الطبيعي المقطوع الأيسر ذو الحد الأدنى $\gamma = 0$. من المهم تطبيق التوزيع الطبيعي المقطوع الأيسر على حسابات المخزون بالنسبة لهذه القطع التي تكون طلباتها الشهرية غير طبيعية. وعندما يستند مقدار المخزون الذي ينبغي توفيره لكل قطعة حسب الموقع على التوزيع الطبيعي، تكون النتائج خاطئة باعتبار أن معظم الطلبات ليست توزيعاً طبيعياً، بل تتبع التوزيع الطبيعي المقطوع الأيسر. ونورد فيما يلي تجربة المؤلف في صناعة السيارات وفي قطاع البيع بالتجزئة.

١٠-٩ مركز التوزيع في قطّاع السيارات

في مركز توزيع كبير يضم ما يزيد عن 100000 قطعة، يتم وضع التوقعات كل شهر وتتم مراقبة قرارات الشراء كل يوم. كما يتم حساب معامل اختلاف الطلبات الشهرية لكل قطعة شهرياً، وقد تم إدراجها في الجدول أدناه من خلال عمودين هما معامل الاختلاف والأجزاء المئوية. إن الأجزاء المئوية هي حصة جميع القطع ذات معامل الاختلاف كما هو مبين من 0 إلى ما يزيد عن 1، كما أن لكل قطعة ذات معامل الاختلاف 0.50 أو أقل طلبات شهرية تشبه التوزيع الطبيعي، وأن لكل قطعة ذات معامل الاختلاف 0.50 أو أكبر طلبات شهرية تتبع التوزيع الطبيعي المقطوع الأيسر. لتقديم أفضل الحسابات عن الزمن المناسب للشراء وكميته، ينبغي استخدام الحسابات التي تستخدم التوزيع الطبيعي المقطوع الأيسر مع القطع ذات معامل الاختلاف 0.50 أو أكبر.

الأجزاء المئوية	معامل الاختلاف
26	0.00–0.30
12	0.30–0.50
12	0.50–0.80
10	0.80–1.00
40	1.00 –
100	المجموع

١٠-١٠ الوكيل، أو بائع التجزئة، أو المتجر

يوجد لدى أحد الوكلاء النموذجيين 20000 قطعة وهو محدود بمقدار المخزون الذي ينبغي الاحتفاظ به نظراً لقيود المساحة والميزانية. إن هدفه هو توفير الحد الأدنى لاستثمار المخزون لتحقيق مستوى خدمة عالٍ. وعندما ينفذ مخزونه، يكون مورّده هو مركز التوزيع. إن الطلبات الشهرية لكل وكيل (متجر، بائع تجزئة) هي أدنى بكثير من الطلبات ذات القطعة في مركز التوزيع. إن الطلبات الشهرية لا تتبع التوزيع الطبيعي، بل توزيع بواسون أو التوزيع الطبيعي المقطوع الأيسر، وليس هناك طلبات شهرية طبيعية بالنسبة لجميع القطع تقريباً ضمن هذا السيناريو، لذا ينبغي إجراء حسابات ضبط المخزون باستخدام التوزيع الطبيعي المقطوع الأيسر.

١٠-١١ ملخص

إنَّ للتوزيع الطبيعي المقطوع الأيسر القياسي معلمة واحدة k وهي التي تشكّل التوزيع الذي يتراوح ما بين طبيعي وشبه أسي. يُشار إلى المتغير القياسي بأنه t وله قيم الصفر أو أكبر. كما تم وضع جدول يبيّن القياسات الإحصائية الرئيسة للمتغير t بالنسبة للقيم المختارة للمعلمة k وهي تتراوح ما بين 3.0− و 3.0+. إن القياسات الإحصائية للمتغير t هي النقط المئوية α عندما يكون $\alpha = 0.01$ و 0.99، والمتوسط، والانحراف المعياري، ومعامل الاختلاف، ونسبة الانتشار. وعند توفر بيانات العينة من المتغير X، تتيح القياسات الإحصائية المأخوذة من البيانات للمحلل تقدير المعلمة k التي تناسب بيانات العينة؛ كما تبيّن كيفية تقدير الحد الأدنى γ حيث إن $\gamma \leq x$، وأي $x\alpha$ حيث إن $P(X \leq x\alpha) = \alpha$.

الفصل الحادي عشر

التوزيع الطبيعي المقطوع الأيمن
RIGHT TRUNCATED NORMAL DISTRIBUTION

١١-١ مقدمة

في عام 2001، قام آرفيد جونسون ونيك ثوموبولوس (Arvid Johnson and Nick Thomopoulos) بوضع جداول عن التوزيع الطبيعي المقطوع الأيمن. يتّخذ التوزيع الطبيعي المقطوع الأيمن عدة أشكال تتراوح ما بين طبيعي وشبه أسي، كما أن للتوزيع معلمة واحدة وهي k حيث يتضمن المدى جميع قيم التوزيع الطبيعي القياسي التي تكون أقل من قيمة $z = k$. فـي هذه الحالة، يتّخذ التوزيع شكل التوزيع الطبيعي القياسي علـى الجانب الأيسر وينقطع على الجانب الأيمن. يُشار إلى المتغير العشوائي بأنه t حيث إن قيمته تكون مساوية للصفر أو قيمة سالبة. عند تحديد k، يتم حساب الإحصاءات التالية: المتوسط، والانحراف المعياري، ومعامل الاختلاف، والنقطتين المئويتين 0.01% و0.99% للمتغير t . كما يتم حساب نسبة الانتشار للتوزيع الطبيعي المقطوع الأيمن بالنسبة لكل قيمة من قيم المعلمة k . لقد تم وضع جدول يبيّن جميع إحصاءات قيم k التي تتراوح ما بين -3.0 و $+3.0$. وعند توفر بيانات العينة، يحسب المحلل الإحصاءات التالية: متوسط العينة، والانحراف المعياري، والحد الأدنى ، والحد الأقصى. وبناءً على ذلك، يتم حساب تقدير نسبة الانتشار، ويُقارن هذا التقدير بالقيم الجدولية لوضع المعلمة k ذات القيمة الأقرب إلى θ . وعند تقدير المعلمة k بالنسبة لبيانات العينة، يـحدد المحلل التوزيع الطبيعي المقطوع الأيمن الذي يعد الأفضل بالنسبة للبيانات. وانطلاقاً من ذلك، يمكن تقدير الحد الأعلـى δ وكذلك أي نقطة مئوية α بالنسبة للمتغير X ، كما يشير اختبار نسبة الانتشار أحياناً إلـى أنه من الأفضل تمثيل بيانات العينة من خلال التوزيع الطبيعي، وفي حالاتٍ أخرى من خلال التوزيع الطبيعي المقطوع الأيمن.

١١-٢ أساسيات

يتيح التوزيع الطبيعي المقطوع الأيمن للمحلل تطبيق مجموعة متنوعة من الأشكال التي تتراوح ما بين طبيعي إلى شبه أسي، كما إن للتوزيع معلمة واحدة k تحدد شكله باعتبار أنه لا يطبّق التوزيع الطبيعي القياسي للقيم التي تكون أقل من $z = k$. نورد فيما يلي استعراضاً سريعاً عن التوزيع الطبيعي القياسي ويلي ذلك وصفاً كاملاً.

١١-٣ التوزيع الطبيعي القياسي

إنَّ للمتغير العشوائي z والذي يمثل التوزيع الطبيعي القياسي حدين هما ∞- و ∞+، المتوسط له صفر والتباين واحد ونكتب:

$$Z \sim N(0,1)$$

نورد فيما يلي دالة الكثافة الاحتمالية للمتغير Z، ودالة الاحتمال التراكمي، والاحتمال المكمّل، كما ننوه بأنه لا توجد صيغة محددة لدالة الاحتمال التراكمية وبالتالي فقد تم وضع العديد من الطرق الكمّية لحساب $F(z)$. تم توضيح إحدى هذه الطرق في الفصل الثامن ويتم استخدامها في بعض الحسابات في الفصل الحالي.

دالة الكثافة الاحتمالية للمتغير z = $f(z) = \left(1\big/\sqrt{2\pi}\right)e^{-z^2/2}$

دالة الاحتمال التراكمية عندما $z = k$ ، $F(k) = P(z \le k) = \int\limits_{-\infty}^{k} f(z)\,dz$

دالة الاحتمال المكمّل عندما $z = k$ $H(k) = P(z > k) = 1 - F(k)$

نبيّن فيما يلي بعض التكاملات ذات الصلة بالتوزيع الطبيعي القياسي:

$$\int\limits_{-\infty}^{k} zf(z)\,dz = -f(k)$$

$$\int\limits_{-\infty}^{k} z^2 f(z)\,dz = -kf(k) + F(k)$$

بعد ذلك يتم التوصل إلى متوسط المتغير t وتباينه وانحرافه المعياري على النحو التالي:

متوسط t عند $k = E(t)_k = \mu_t(k)$

تباين t عند $k = E(t^2)_k - E(t)_k^2 = V(t)_k$

الانحراف المعياري للمتغير t عند k $\sigma_t(k) = \sqrt{V(t)_k}$

علـى الرغم من أن مدى z هو $(-\infty,+\infty)$، فإن جميع الاحتمالات تقريباً تقع ما بين -3.0 و $+3.0$. ولتبسيط حساب القياس الإحصائي في التوزيع الطبيعي المقطوع الأيمن، يتم استخدام المدى ± 3 بالنسبة للمتغير Z . من خلال هذا المدى المعدّل، يصبح الحدين الأعلى والأدنى للمتغير t على النحو التالي: $\left[0, -(3+k)\right]$. ويلاحظ أنه عندما تكون $k = 3$ ، يكون مدى t من (0 إلى 6-)، وعندما تكون $k = 2$ ، يكون مدى t من (0 إلى 5−) وهكذا.

١١-٥ الاحتمال التراكمي للمعلمة K

تمثّل النقطة المئوية α للمتغير t قيمة t ، حيث إن احتمال أن يكون t أقل من أو يساوي $t\alpha$ هو α ، أي أن

$$P(t \le t\alpha) = \alpha$$

ذلك مثل الاحتمال التراكمي للمقدار $t\alpha$ كما هو مبيّن أدناه:

$$G(t\alpha) = \alpha$$

تبيّن الخطوات الأربع التالية كيفية إيجاد $t\alpha$ ، وقيمة النقطة المئوية α للمتغير t بالنسبة لمجموعتي k و $G(t)$

١. بالنسبة لقيم معلومة للمعلمة k ، يبيّن الجدول (٨-١) قيم $F(k)$.

٢. باعتبار أن $G(t) = \alpha = F(z)\big/F(k)$:

$$F(z) = \alpha F(k)$$

٣. من خلال $F(z)$ ، يبيّن الجدول (٨-١) قيمة z المقابلة.

٤. و أخيراً

$$t\alpha = (z - k)$$

عندما يكون $\alpha = 0.01$ ، و 0.99 ، يتم التوصل إلى النقطتين المئويتين التاليتين:

$t_{0.01} =$ النقطة المئوية 0.01

$t_{0.99} =$ النقطة المئوية 0.99

١١-٦ متوسط المتغير t وانحرافه المعياري

إنَّ متوسط المتغير t وانحرافه المعياري يعرفان على النحو التالي:

$\mu_t(k) =$ متوسط المتغير t ذي المعلمة k

$\sigma_t(k) =$ الانحراف المعياري للمتغير t ذي المعلمة k

١١-٧ نسبة انتشار التوزيع الطبيعي المقطوع الأيمن

إن نسبة الانتشار المُشار إليها بالرمز θ تكون خاصة بكل معلمة k ويتم حسابها كما هو مبيّن أدناه:

$$\theta = \left[\mu_{t(k)} - t_{0.01}\right] / \left[t_{0.99} - \mu_{t(k)}\right] = \text{نسبة الانتشار}$$

١١-٨ القيم الجدولية

يبيّن الجدول (١١-١) قائمة القياسات الرئيسة في التوزيع الطبيعي المقطوع الأيمن حيث يكون مدى k على النحو التالي: $[-3.0,(0.1),+3.0]$. بالنسبة لكل قيمة مدخلة للمعلمة k، تكون القياسات المدرجة على النحو التالي: النقطتان المئويتان α $t_{0.01}, t_{0.99}$، والمتوسط، والانحراف المعياري، ومعامل الاختلاف، ونسبة الانتشار.

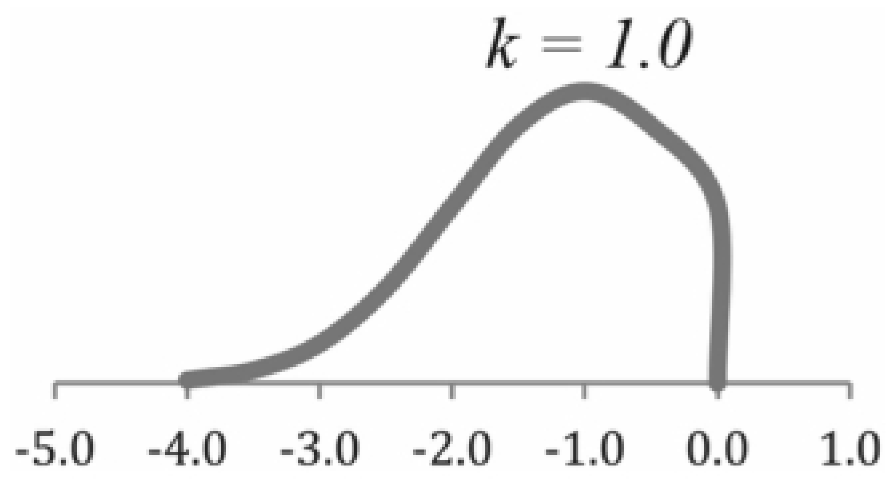
جدول (١١-١). التوزيع الطبيعي المقطوع الأيمن من خلال المعلمة k، والنقطتين المئويتين $t_{0.01}, t_{0.99}$، والمتوسط، $\mu_t(k)$، والانحراف المعياري $\sigma_t(k)$، ومعامل الاختلاف، $\mathrm{cov}(k)$، ونسبة الانتشار θ.

k	$t_{0.01}$	$t_{0.99}$	$\mu_{t(k)}$	$\sigma_{t(k)}$	$\mathrm{cov}(k)$	θ
-3.0	-1.494	-0.006	-0.283	0.264	-0.93	4.36
-2.9	-1.397	-0.005	-0.291	0.270	-0.93	3.88
-2.8	-1.365	-0.005	-0.298	0.277	-0.93	3.64
-2.7	-1.359	-0.005	-0.306	0.284	-0.93	3.50
-2.6	-1.366	-0.005	-0.314	0.291	-0.93	3.40

تابع جدول (١١-١).

k	$t_{0.01}$	$t_{0.99}$	$\mu_{t(k)}$	$\sigma_{t(k)}$	$\mathrm{cov}(k)$	θ
0.7	-3.129	-0.024	-1.112	0.736	-0.66	1.86
0.8	-3.215	-0.027	-1.168	0.756	-0.65	1.80
0.9	-3.303	-0.030	-1.226	0.775	-0.63	1.74
1.0	-3.391	-0.034	-1.288	0.794	-0.62	1.68
1.1	-3.481	-0.039	-1.352	0.812	-0.60	1.62
1.2	-3.573	-0.044	-1.419	0.830	-0.58	1.57
1.3	-3.665	-0.051	-1.490	0.847	-0.57	1.51
1.4	-3.759	-0.059	-1.563	0.863	-0.55	1.46
1.5	-3.853	-0.068	-1.639	0.879	-0.54	1.41
1.6	-3.948	-0.080	-1.717	0.894	-0.52	1.36
1.7	-4.044	-0.094	-1.798	0.907	-0.50	1.32
1.8	-4.141	-0.110	-1.882	0.920	-0.49	1.28
1.9	-4.238	-0.131	-1.968	0.931	-0.47	1.24
2.0	-4.336	-0.155	-2.055	0.942	-0.46	1.20
2.1	-4.434	-0.184	-2.145	0.951	-0.44	1.17
2.2	-4.532	-0.218	-2.236	0.959	-0.43	1.14
2.3	-4.631	-0.259	-2.329	0.966	-0.41	1.11
2.4	-4.730	-0.305	-2.423	0.972	-0.40	1.09
2.5	-4.829	-0.359	-2.518	0.978	-0.39	1.07
2.6	-4.929	-0.419	-2.614	0.982	-0.38	1.06
2.7	-5.028	-0.486	-2.710	0.986	-0.36	1.04
2.8	-5.128	-0.559	-2.808	0.989	-0.35	1.03
2.9	-5.228	-0.638	-2.906	0.991	-0.34	1.02
3.0	-5.328	-0.721	-3.004	0.993	-0.33	1.02

كقاعدة عامة عندما يكون تقدير بيانات العينة θ مساوياً 1.30 أو أقل، لا يختلف التوزيع الطبيعي المقطوع الأيمن بشكلٍ كبير عن التوزيع الطبيعي. وبالتالي يُطبّق التوزيع الطبيعي، وإذا لم يتحقق ذلك يعد التوزيع الطبيعي المقطوع الأيمن أفضل مع هذه الحالة.



شكل (١١-١). التوزيع الطبيعي المقطوع الأيمن عندما $k = 1.00$.

١١-٩ بيانات العيّنة

عند توفر بيانات العينة $(x_1, \ldots, x_n)$، يمكن للمحلل التوصل إلى القياسات الإحصائية التالية فيما يتعلق بالمتغير X:

$\overline{x} =$ المتوسط

$s =$ الانحراف المعياري

$x(1) =$ الحد الأدنى $(x_1, \ldots, x_n)$

$x(n) =$ الحد الأعلى $(x_1, \ldots, x_n)$

بعد جمع القياسات الإحصائية، يتم حساب تقدير نسبة الانتشار للتوزيع الطبيعي المقطوع الأيمن المُشار إليها بالرمز $\hat{\theta}$ على النحو التالي:

$$\hat{\theta} = \left[\overline{x} - x(1)\right] / \left[x(n) - \overline{x}\right]$$

نلاحظ أنه عند استخدام التوزيع الطبيعي المقطوع الأيمن، يكون $\hat{\theta}$ أكبر من 1، وعند تقدير نسبة الانتشار، يبحث المحلل في الجدول (١١-١) لإيجاد المعلمة k من خلال أقرب θ، ومن ثم التوصل إلى القياسات التالية للتوزيع الطبيعي المقطوع الأيمن:

$k =$ معلمة التوزيع الطبيعي المقطوع الأيمن

$\mu_t(k) =$ متوسط t ذي المعلمة k

$\sigma_t(k) =$ الانحراف المعياري للمتغير t ذي المعلمة k

$t_{0.99} =$ النقطة المئوية 0.99 للمتغير t ذي المعلمة k

الفصل الثاني عشر

التوزيع المثلثي
TRIANGULAR DISTRIBUTION

١٢-١ مقدمة

عندما يحتاج المحلل إلى التوزيع المتصل في أحد المشاريع و ليس لديه إلا بعض المعلومات عن الشكل، فإنه غالباً ما يستخدم التوزيع المثلثي و يسعى إلى تقدير القيم الدنيا، والعليا، والمرجّحة للمتغير ومن هذه القيم يشكّل التوزيع. وترتفع دالة الكثافة الاحتمالية بشكلٍ خطي من القيمة الدنيا إلى المنوال، وتنخفض بشكلٍ خطي نحو القيمة العليا. أما التوزيع المتصل الآخر فهو التوزيع المثلثي القياسي الذي يتراوح ما بين الصفر والواحد. إن التوزيع الأخير أسهل للإدارة من الناحية الرياضية، ويُستخدم لتحديد التوزيع الاحتمالي، والاحتمال التراكمي، والمتوسط، والتباين، والانحراف المعياري. ويمكن تحويل جميع الإحصاءات والاحتمالات من التوزيع المثلثي القياسي إلى التوزيع المثلثي بسهولة.

١٢-٢ أساسيات

هناك نوعان للتوزيع المثلثي هما: التوزيع المثلثي القياسي ذو المتغير العشوائي Y، والتوزيع المثلثي ذو المتغير العشوائي X. إن التوزيع المثلثي هو التوزيع الذي يستخدمه المحلل في التطبيق البحثي، في حين أن التوزيع المثلثي القياسي أسهل للتطبيق من الناحية الرياضية. ونبيّن فيما يلي كلا التوزيعين ونوضّح كيفية تحويل أحدهما للآخر.

١٢-٣ التوزيع المثلثي القياسي

يتراوح مدى المتغير العشوائي للتوزيع المثلثي القياسي Y ما بين صفر وواحد، ويُشار إلى القيمة المرجّحة بالرمز $\tilde{y}$، كما هو ملاحظ أدناه:

$0 \le y \le 1$

$0 =$ القيمة الدنيا للمتغير Y

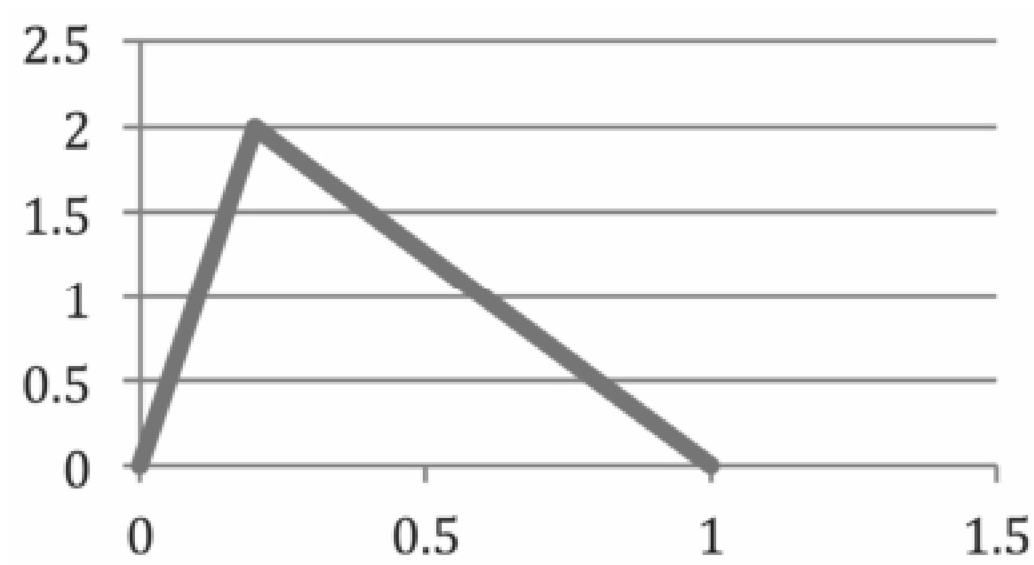
$1 =$ القيمة العليا للمتغير Y

$\tilde{y} =$ منوال المتغير Y

ونرمز لذلك اختصاراً بالرمز:

$$Y \sim T(0,1,\tilde{y})$$

يمثل الشكل (١٢-١) دالة كثافة التوزيع المثلثي القياسي.



شكل (١٢-١). التوزيع المثلثي القياسي $T(0,1,\tilde{y})$.

نبيّن فيما يلي احتمال المتغير Y الذي يرتفع بشكل خطي من 0 إلى $\tilde{y}$، ثم ينخفض من $\tilde{y}$ إلى 1.

$$f(y) = 2y/\tilde{y} \qquad y = (\text{من 0 إلى } \tilde{y})$$

$$= 2(1-y)/(1-\tilde{y}) \qquad y = (\text{من } \tilde{y} \text{ إلى 1})$$

كما أن دالة الاحتمال التراكمي للمتغير Y تكون على النحو التالي:

$$F(y) = y^2/\tilde{y} \qquad y = (\text{من 0 إلى } \tilde{y})$$

$$= 1-(1-y)^2/(1-\tilde{y}) \qquad y = (\text{من } \tilde{y} \text{ إلى 1})$$

نورد فيما يلي المتوسط، والتباين، والانحراف المعياري للمتغير Y:

$$\mu_y = (1+\tilde{y})/3$$
$$\sigma_y^2 = (1+\tilde{y}^2-\tilde{y})/18$$
$$\sigma_y = \sqrt{\sigma_y^2}$$

١٢-٤ التوزيع المثلثي

يُشار إلى المتغير المُقابل للمتغير Y بالمتغير X، ويُشتق من التوزيع المثلثي. يتراوح مدى المتغير X ما بين a و b، ويُشار إلى القيمة المرجّحة بـ $\tilde{x}$ كما هو مبيّن أدناه:

$a =$ القيمة الدنيا للمتغير X

$b =$ القيمة العليا للمتغير X

$\tilde{x} =$ منوال المتغير X

وبالتالي، باستخدام المعالم: a، b، $\tilde{x}$ نكتب اختصاراً:

$$X \sim T(a,b,\tilde{x})$$

ويكون التحويل من المتغير Y إلى المتغير X، وكذلك من X إلى Y كما يلي:

$$x = a + y(b-a)$$
$$y = \frac{(x-a)}{(b-a)}$$

نبيّن فيما يلي كيفية تحويل متوسط المتغير X، وتباينه، وانحرافه المعياري مما يقابله في Y، وكذلك كيفية التوصل إليها من المعالم a، b و $\tilde{x}$.

$$\begin{aligned}\mu_x &= a + \mu_y\left(b-a\right)\\ &= \left(a+b+\tilde{x}\right)/3\\ \sigma_x^2 &= \sigma_y^2\left(b-a\right)^2\\ &= \left(a^2+b^2+\tilde{x}^2-ab-a\tilde{x}-b\tilde{x}\right)/18\\ \sigma_y &= \sqrt{\sigma_x^2}\end{aligned}$$

المثال (١٢-١). على افتراض أن X هو متغير التوزيع المثلثي ذو المعالم: $X \sim (10,90,70)$ يتم حساب متوسط X وتباينه وانحرافه المعياري على النحو المبين أدناه.

$$\mu_x = (10+90+70)\big/3 = 56.67$$
$$\sigma_x^2 = (10^2+90^2+70^2-10\times 90-10\times 70-90\times 70)\big/18 = 289$$
$$\sigma_x = \sqrt{289} = 17.0$$

ونوضّح فيما يلي كيفية التحويل إلى التوزيع المثلثي القياسي

$$\tilde{y} = (70-10)\big/(90-10) = 0.75$$

$$Y \sim T(0,1,0.75)$$

وتكون دالة الكثافة الاحتمالية والاحتمال التراكمي للمتغير Y على النحو التالي:

$f(y) = 2y\big/0.75 = y$ (من 0 إلى 0.75)

$= 2(1-y)\big/0.25 = y$ (من 0.75 إلى 1.00)

$F(y) = y^2\big/0.75 = y$ (من 0 إلى 0.75)

$= 1-(1-y)^2\big/0.25 = y$ (من 0.75 إلى 1.00)

يتم التوصل إلى متوسط المتغير Y وتباينه وانحرافه المعياري كما هو مبين أدناه:

$$\mu_y = (1+0.75)\big/3 = 0.583$$
$$\sigma_y^2 = (1+0.75^2-0.75)\big/18 = 0.045$$
$$\sigma_y = \sqrt{0.045} = 0.212$$

ومن ثم يتم التوصل إلى متوسط المتغير X وتباينه وانحرافه المعياري ومنواله كما هو مبين أدناه:

$$\mu_x = 10+0.583(90-10) = 56.67$$
$$\sigma_x^2 = 0.045(90-10)^2 = 289$$
$$\sigma_x = \sqrt{289} = 17$$
$$\tilde{x} = 10+0.75(90-10) = 70$$

١٢-٥ القيم الجدولية للمتغير Y

يبيّن الجدول (١٢-١) قيم الاحتمال التراكمي $F(y)$ بالنسبة للقيم المختارة للمتغير y والمنوال $\tilde{y}$.

المثال (١٢-٢). عندما يكون $\tilde{y} = 0.3$ ، يبين الجدول (١٢-١) أن

$$F(0.2) = 0.13,\ F(0.8) = 0.94,\ \mu_y = 0.43 \text{ و } \sigma_y = 0.21$$

و فيما يلي نوضح كيفية حساب هذه القيم:

جدول (١٢-١). الاحتمال التراكمي $F(y)$ للتوزيع المثلثي القياسي بالنسبة للقيم المختارة للمتغير y ، والقيم المختارة للمنوال $\tilde{y}$.

$y/\tilde{y}$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.1	0.19	0.10	0.05	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01
0.2	0.36	0.29	0.20	0.13	0.10	0.08	0.07	0.06	0.05	0.04	0.04
0.3	0.51	0.46	0.39	0.30	0.22	0.18	0.15	0.13	0.11	0.10	0.09
0.4	0.64	0.60	0.55	0.49	0.40	0.32	0.27	0.23	0.20	0.18	0.16
0.5	0.75	0.72	0.69	0.64	0.58	0.50	0.42	0.36	0.31	0.28	0.25
0.6	0.84	0.82	0.80	0.77	0.73	0.68	0.60	0.51	0.45	0.40	0.36
0.7	0.91	0.90	0.89	0.87	0.85	0.82	0.77	0.70	0.61	0.54	0.49
0.8	0.96	0.96	0.95	0.94	0.93	0.92	0.90	0.87	0.80	0.71	0.64
0.9	0.99	0.99	0.99	0.99	0.98	0.98	0.97	0.97	0.95	0.90	0.81
1.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
μ	0.33	0.37	0.40	0.43	0.47	0.50	0.53	0.57	0.60	0.63	0.67
σ	0.24	0.22	0.22	0.21	0.21	0.20	0.21	0.21	0.22	0.22	0.24

كما تم إدراج متوسط المتغير Y وانحرافه المعياري بالنسبة لكل منوال.

$$F(0.2) = 0.2^2 / 0.3 = 0.133$$
$$F(0.8) = 1 - (1 - 0.8^2) / 0.7 = 0.942$$

$$\mu_y = (1+0.3)/3 = 0.433$$
$$\sigma_y^2 = (1+0.3^2 - 0.3)/18 = 0.044$$
$$\sigma_y = 0.044 = 0.209$$

١٢-٦ اشتقاق $x\alpha =$ النقطة المئوية $- \alpha$ للمتغير X

لإيجاد $x\alpha$ =النقطة المئوية α للمتغير x، يتم اتخاذ الخطوات الثلاث التالية.

١. يتم التوصل إلى المنوال $\tilde{y}$ للمتغير Y من معالم المتغير X كما هو مبيّن أدناه:

$$\tilde{y} = (\tilde{x} - a)/(b-a)$$

٢. يتم حساب النقطة المئوية α للمتغير Y وهي $y\alpha$ على النحو التالي بناءً على العلاقة بين α و $\tilde{y}$.

إذا كان $\alpha \le \tilde{y}$

$$\alpha = y\alpha^2/\tilde{y}$$
$$y\alpha = \sqrt{\alpha\tilde{y}}$$

وإذا كان $\alpha > \tilde{y}$

$$\alpha = 1-(1-y\alpha)^2/(1-\tilde{y})\frac{n!}{r!(n-r)!}$$
$$y\alpha = 1-\sqrt{(1-\alpha)(1-\tilde{y})}$$

٣. يتم التوصل إلى النقطة المئوية α للمتغير X كما هو مبين أدناه:

$$x\alpha = a + y\alpha(b-a)$$

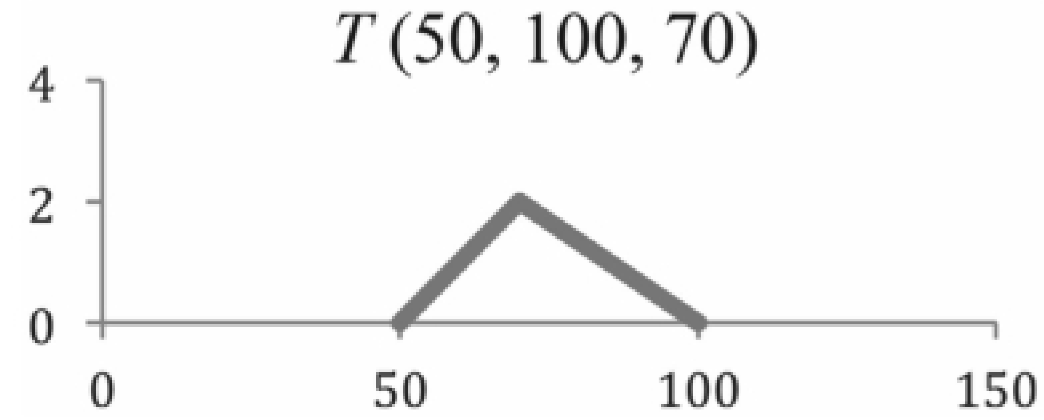
المثال (١٢-٢). بالنظر في التوزيع المثلثي حيث إن $X \sim (50,100,70)$، الموضح بالتمثيل البياني دالة الكثافة الاحتمالية في الشكل (١٢-٢).

نلاحظ أن منوال التوزيع المثلثي القياسي يعرف على النحو التالي:

$$\tilde{y} = (70-50)/(100-50) = 0.40$$

وبالتالي

$$Y \sim T(0,1,0.40)$$



شكل (١٢-٢). التوزيع المثلثي ذو المعالم (50، 100، 70).

على افتراض أن المحلل يسعى إلى تقدير $x_{0.10}$ و $x_{0.90}$. حيث تم الحصول عليهما كما هو مبين أدناه:

حيث إن $0.10 \leq 0.40$:

$$y_{0.10} = \sqrt{0.10 \times 0.40} = 0.20$$
$$x_{0.10} = 50 + 0.20(100 - 50) = 60.00$$

و حيث إن $0.90 \leq 0.40$:

$$y_{0.90} = 1 - \sqrt{\left(1 - 0.90\right)\left(1 - 0.40\right)} = 0.755$$
$$x_{0.90} = 50 + 0.755(100 - 50) = 87.75$$

١٢-٧ تقدير المعالم عند عدم توفر بيانات

يتم استخدام هذا التوزيع بشكلٍ رئيس عندما لا يتوفر لدى المحلل أي بيانات حيث يتم توفير تقدير عن القيمة الدنيا، والعليا، والمرجّحة للمتغير X ، وينتج عن ذلك ما يلي:

$a =$ القيمة الدنيا

$b =$ القيمة العليا

$\tilde{x} =$ القيمة المرجّحة

وبالتالي

$$X \sim T(a, b, \tilde{x})$$

١٢-٨ ملخص

إنَّ للتوزيع المثلثي القياسي المتغير العشوائي Y ويتراوح مداه ما بين الصفر وواحد. يرتفع التوزيع بشكل خطي من الصفر إلى المنوال ثم ينخفض من المنوال إلى واحد. تم توضيح دالة الكثافة الاحتمالية، والاحتمال التراكمي، والمتوسط، والانحراف المعياري للتوزيع المثلثي القياسي، كما يتراوح مدى التوزيع المثلثي ذا المتغير X ما بين a و b ويكون المنوال بينهما. يمكن التحويل بسهولة من المتغير Y إلى X، وكذلك من X إلى Y. ويتم استخدام التوزيع عندما يحتاج المحلل إلى تطبيق التوزيع في أحد المشاريع ولا يكون لديه سوى تقدير عن القيمة الدنيا، والعليا، والمرجّحة للمتغير X. وعند الضرورة يتم اشتقاق النقطة المئوية α للمتغير X بسهولة من النقطة المئوية α للمتغير Y التي تم التوصل إليها أولاً.

الفصل الثالث عشر

التوزيع المنتظم المنفصل
DISCRETE UNIFORM DISTRIBUTION

١٣-١ مقدمة

يحدث التوزيع المنتظم المنفصل عندما يمكن للمتغير العشوائي X أن يأخذ أي قيمة عدد صحيح من a إلى b ذات احتمال متساوٍ. إنَّ اختيار أي رقم من 1 إلى 1000 هو عبارة عن نظام من المرجّح أن يكون فيه الرقم الرابح x أي رقم من الأرقام الموجودة ضمن النطاق المقبول. غالباً ما تكون المعلمتان (a و b) معلومتين للمحلل الذي يسعى إلى تطبيق التوزيع. وفي أحيانٍ أخرى، تكون المعلمتان مجهولتين ويتم توفير بيانات العينة لتقدير المعلمتين. إلا أنه في بعض الأحيان عندما تكون المعلمتان مجهولتين وبيانات العينة غير متاحة، يطلب المحلل مشورة أحد الخبراء الذي يقدم بعض التقريبات عن المدى، وتُستخدم هذه المعلومات لتقدير قيم المعلمتين.

١٣-٢ أساسيات

عند تطبيق التوزيع المنتظم المنفصل، يمكن للمتغير X أن يأخذ أي عدد صحيح من a إلى b، حيث إن a هو القيمة الدنيا و b هو القيمة العليا. وفي هذه الحالة تكون (a و b) معلمتي المتغير X حيث يكون المدى:

$$x = a, a+1, \ldots, b$$

باعتبار أنه من المرجح أن تحدث القيم المقبولة للمتغير X، تكون احتمالية أي قيمة من قيم x على النحو التالي:

$$P(x) = 1\big/(b - a + 1) \qquad x = a, a+1, \dots, b$$

ويصبح الاحتمال التراكمي للمتغير X أو أصغر على النحو التالي:

$$F(x) = P(w \le x) = (x - a + 1)\big/(b - a + 1) \qquad x = a, a+1, \dots, b$$

كما يتم التوصل إلى متوسط وتباين X على النحو المبين أدناه:

$$\mu = (a + b)\big/2$$
$$\sigma^2 = \left[(b - a + 1)^2 - 1\right]\big/12$$

المثال (١٣-١). خلال فترة حصاد التفاح، يتم جني التفاح ووضعه في السلة إلى أن تمتلئ. يختلف عدد التفاح في السلة بناءً على حجم التفاح، ويتراوح المدى ما بين 50 و60 تفاحة. وعند تمثيل المتغير X لعدد التفاح في السلة كمتغير عشوائي، يكون المدى ما بين 50 و60 ، ويكون احتمال أي قيمة من قيم X على النحو التالي:

$P(x) = 1\big/\left(60 - 50 + 1\right) = 0.091$ و x تتراوح بين 50 و 60

كما يتم حساب الاحتمال التراكمي لوضع 53 تفاحة في السلة أو أقل على النحو التالي:

$$P(X \le 53) = (53 - 50 + 1)\big/60 - 50 + 1 = 0.364$$

ويتم التوصل إلى متوسط المتغير X ، وتباينه، وانحرافه المعياري على النحو المبين أدناه:

$$\mu = \left(50 + 60\right)\big/2 = 55$$
$$\sigma^2 = \left[\left(60 - 50 - 1\right)^2 - 1\right]\big/12 = 10$$
$$\sigma = 3.16$$

١٣-٣ نسبة لكسيس

عندما يكون $a = 0$ ، تصبح نسبة لكسيس على النحو التالي:

$$\begin{aligned}\tau &= \sigma^2 \big/ \mu \\ &= \left[(b+1)^2 - 1\right] \big/ 6b\end{aligned}$$

مع ملاحظة $b \geq 4,\ \tau \geq 1$.

١٣-٤ بيانات العيّنة

عند توفر عيّنة $(x_1, \ldots, x_n)$ حجمها n، يتم التوصل إلى الإحصاءات التالية:

$\tilde{x} =$ متوسط العينة

$s^2 =$ تباين العينة

$x(1) =$ القيمة الدنيا $(x_1, \ldots, x_n)$

$x(n) =$ القيمة العليا $(x_1, \ldots, x_n)$

١٣-٥ تقدير المعلمتين عند توفر بيانات العيّنة

يتم استخدام الإحصاءات أعلاه من بيانات العينة لتقدير المعلمتين كما هو موضّح أدناه. ينتج عن طريقة مقدّر الإمكان الأعظم ما يلي:

$$\begin{aligned}\hat{a} &= x(1) \\ \hat{b} &= x(n)\end{aligned}$$

وتتمثل الطريقة الأخرى لتقدير المعالم في طريقة العزوم كما هو موضّح أدناه:

$\hat{a} =$ العدد الصحيح الأدنى للمقدار $\left(\overline{x} + 0.5 - 0.5\sqrt{12s^2 + 1}\right)$

$\hat{b} =$ العدد الصحيح الأعلى للمقدار $\left(\overline{x} + 0.5 - 0.5\sqrt{12s^2 + 1}\right)$

المثال (١٣-٢). على افتراض أن البيانات تمثل أعداداً صحيحة منفصلة وإحدى التجارب تعطي المدخلات العشر للعينة كالتالي

(5,7,9,8,14,16,8,4,11,13)

وهناك حاجة إلى تقدير المعلمتين (a، b). عليه يتم حساب الإحصاءات التالية من البيانات:

$$\overline{x} = 9.5$$
$$s^2 = 15.39$$
$$x(1) = 4$$
$$x(n) = 16$$

وباستخدام طريقة مقدّر الإمكان الأعظم، يكون تقدير المعلمتين على النحو التالي:

$$\hat{a} = 4$$
$$\hat{b} = 16$$

وكذلك يكون تقدير المعلمتين باستخدام طريقة العزوم على النحو التالي:

$\hat{a} = \left(9.5 + 0.5 - 0.5\sqrt{12 \times 15.39 + 1}\right) =$ القيمة الدنيا $(3.19) = 3$

$\hat{b} = \left(9.5 - 0.5 + 0.5\sqrt{12 \times 15.39 + 1}\right) =$ القيمة العليا $(15.81) = 16$

المثال (١٣-٣). تشعر قيادة الاستخبارات العسكرية بالقلق حيال عدد وحدات إحدى الأسلحة التدميرية التي أنتجها العدو. وخلال مناوشاتهم، تم استعادة الوحدات n لهذا السلاح وقد حُفر عليها أرقام تسلسلية. تم إدراج الأرقام المُلتقطة على شكل: $(x_1, \ldots, x_n)$، ومن خلال الأرقام، تم التوصل إلى الإحصاءات العادية على شكل: $\tilde{x}$، s، $x(1), x(n)$. ومن خلال هذه البيانات، والطريقتين اللتين تم توضيحهما سابقاً، يتم حساب تقدير المعلمتين $\left(\hat{b}, \hat{a}\right)$. على افتراض أنه قد تم إدراج الأرقام التسلسلية في ترتيب صحيح، يتم تقدير عدد أسلحة العدو التي تندرج ضمن هذا النوع على النحو التالي:

$$\mathbb{N} = (\hat{b} - \hat{a} + 1)$$

١٣-٦ تقدير المعلمتين عند عدم توفر بيانات

عند عدم توفر بيانات العينة لتقدير معلمتي التوزيع المنتظم المنفصل، قد يطلب المحلل مشورة أحد الخبراء فيما يتعلق بالتوزيع. إنَّ إحدى الطرق هي عندما يقدم الخبير النوع التالي لتقدير انتشار التوزيع:

$x_1 =$ النقطة المئوية α_1 للمتغير X

$x_2 =$ النقطة المئوية α_2 للمتغير X

حيث إن:

$$\alpha_1 = (x_1 - a + 1)\big/(b - a + 1)$$
$$\alpha_2 = (x_2 - a + 1)\big/(b - a + 1)$$

من خلال هذه البيانات والعلاقات وباستخدام بعض العمليات الجبرية، يتم التوصل إلى تقدير المعالم على النحو التالي:

$$a' = \left[\alpha_2\left(x_1 + 1\right) - \alpha_1\left(x_2 + 1\right)\big/[\left(\alpha_2 - \alpha_1\right)\right] b' = \left[x_1 + a\left(\alpha_1 - 1\right) - \left(\alpha_1 - 1\right)\right]\big/\alpha_1$$

$\hat{a} = (a')$ القيمة الدنيا

$\hat{b} = (b')$ القيمة العليا

المثال (١٣-٤). يحتاج أحد المحللين إلى استخدام التوزيع المنتظم المنفصل إلا أنه لا يوجد لديه بيانات العيّنة لتقدير قيم المعالم. ثم يطلب مشورة أحد الخبراء الذي يعطيه التقريبات التالية فيما يتعلق بالتوزيع:

يتراوح منتصف المدى المئوي 80 ما بين 5 و20، والذي يُفسّر على النحو التالي:

$$P(5 \le X \le 20) = 0.80$$

وبالتالي:

$$\alpha_1 = 0.10 \text{ و } x_1 = 5$$

$$\alpha_2 = 0.90 \text{ و } x_2 = 20$$

يتم تقدير قيم المعالم على النحو المبين أدناه:

$$4.125 = \left[0.9(5 + 1) - 0.1(20 + 1)\right]\big/\left[0.9 - 0.1\right] = a'$$

$$21.875 = \left[5 + 4.125(0.1 - 1) - (0.1 - 1)\right]\big/\left[0.1\right] = b'$$

$\hat{a} =$ القيمة الدنيا $(4.125) = 4$

$\hat{b} =$ القيمة العليا $(21.875) = 22$

١٣-٧ ملخص

إنَّ للتوزيع المنتظم المنفصل معلمتين هما (a و b)، حيث يتضمن المتغير العشوائي جميع الأعداد الصحيحة من a إلى b ذات الاحتمال المتساوي. يبيّن الفصل كيفية حساب احتمال المتغير X ، والاحتمال التراكمي للمتغير X ، والمتوسط، والانحراف المعياري. وعندما تكون قيم المعلمتين مجهولة وتتوفر بيانات العينة، يتم استخدام بيانات العينة لتقدير قيم المعلمتين. وعند عدم توفر بيانات العينة، يقدم أحد الخبراء تقريبات عن توزيع المتغير بحيث يمكن استخدام هذه المعلومات لتقدير قيم المعلمتين.

الفصل الرابع عشر

توزيع ثنائي الحدين
BINOMIAL DISTRIBUTION

١٤-١ مقدمة

ينسب بعض المؤرخين أول استخدام لتوزيع ثنائي الحدين إلى جاكوب برنولي (Jakob Bernoulli) وهو عالم رياضيات سويسري بارز في القرن السابع عشر. يُطبّق توزيع ثنائي الحدين عند القيام بعدة محاولات لإحدى التجارب مع ملاحظة نتيجتين فقط لكل محاولة، النجاح والفشل، واحتمال أن تبقى النتائج ذاتها في جميع المحاولات ثابت. يحدث ذلك على سبيل المثال عند رمي حجري النرد خمس مرات ونجاح كل رمية عندما يكون عدد النقاط الظاهرة على السطح العلوي يساوي اثنين. يبقى احتمال نجاح كل رمية ثابتاً، وعدد المحاولات خمسة، واحتمال نجاح كل محاولة هو $p = 1/36$. حيث المتغير العشوائي X في هذا السيناريو والذي يمثل عدد النجاحات في المحاولات الخمس هو: $x = 0,1,2,3,4,5$. يبيّن الفصل احتمال توزيع المتغير العشوائي X ، كما تم إدراج متوسط X ، وتباينه، وانحرافه المعياري، ومنواله. وعندما تكون p مجهولة، يمكن تقديرها باستخدام بيانات العينة، وحتى عند عدم توفر بيانات العينة، يمكن التوصل إلى تقدير p .

١٤-٢ أساسيات

تكون معالم توزيع ثنائي الحدين على النحو التالي:

$n =$ عدد المحاولات

$p =$ احتمال نجاح كل محاولة

عندما تكون محاولات n ذات احتمال p نجاح في كل محاولة، يصبح احتمال نجاحات المتغير X على النحو التالي:

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, \ldots, n$$

ويُشار إلى الاحتمال التراكمي للمتغير X أو أصغر بالدالة $F(x)$ وعندما يكون $x = x_0$ فإن

$$F(x_0) = P(X \leq x_0)$$

ويتم حساب متوسط المتغير X، وتباينه، وانحرافه المعياري على النحو المبين أدناه:

$$\begin{aligned} \mu &= np \\ \sigma^2 &= np(1-p) \\ \sigma &= \sqrt{np(1-p)} \end{aligned}$$

كما يتم اشتقاق منوال المتغير $\tilde{\mu}, X$ على النحو التالي:

إذا لم يكن $p(n+1)$ عدداً صحيحاً: $\tilde{\mu} =$ الحد الأدنى لـ $\left[p(n+1)\right]$

إذا كان $p(n+1)$ عدداً صحيحاً: $\tilde{\mu} = p(n+1) - 1$ و $p(n+1)$

١٤-٣ نسبة لكسيس

إنَّ نسبة لكسيس بالنسبة لتوزيع ثنائي الحدين تكون على النحو التالي:

$$\begin{aligned} \tau &= \sigma^2 / \mu \\ &= np(1-p) / np \\ &= (1-p) \end{aligned}$$

مع ملاحظة أن τ دائماً أقل من واحد.

المثال (١٤-١). تتبع إحدى التجارب ذات المحاولات $n = 10$ محاولات توزيع ثنائي الحدين مع الاحتمال $p = 0.3$. يتم التوصل إلى منوال المتغير X (عدد النجاحات) على النحو المبيّن أدناه:

باعتبار أنَّ $p(n+1) = 0.3(11) = 3.3$ ليس عدداً صحيحاً:

$\tilde{\mu}=$ الحد الأدنى للمقدار $\left[p(n+1)\right]$ = الحد الأدنى للمقدار $0.3(11)\ =$ الحد الأدنى $\left[3.3\right]=3$

٤-١٤ التقريب الطبيعي

يتم تقريب المتغير العشوائي X من خلال التوزيع الطبيعي عندما تكون n كبيرة و $p \leq 0.5$ ويكون $np > 5$

أو عندما $p > 0.5$ مع $n(1-p) > 5$

ويكون متوسط المتغير X ، وتباينه، وانحرافه المعياري، كما هو مبين أدناه:

$\mu = np =$ متوسط المتغير X

$\sigma^2 = np(1-p) =$ تباين المتغير X

$\sigma = \sqrt{np(1-p)} =$ الانحراف المعياري للمتغير X

وبالتالي يصبح التقريب الطبيعي في الصورة:

$$X \sim N(\mu, \sigma^2)$$

يتم تقريب احتمال $x = x_0$ أو أقل من خلال تحويل العدد الصحيح للمتغير X إلى متغير التوزيع الطبيعي القياسي المتصل Z كما هو مبين أدناه:

$$z_0 = \left(x_0 + 0.5 - np\right)\Big/\sqrt{np(1-p)}$$

وبالتالي

$$\begin{aligned} P(X \leq x_0) &\approx P(X \leq x_0 + 0.5) \\ &= F(z_0) \end{aligned}$$

حيث إن $F(z)$ هي دالة الاحتمال التراكمية كما يوضح الجدول (٨-١).

كما يتم التوصل إلى احتمال أن $x = x_0$ من خلال تحويل قيمة العدد الصحيح للمتغير X إلى قيمتين من قيم التوزيع الطبيعي القياسي المتصل كما هو مبين أدناه:

$$z_L = \left(x_0 + .5 - np\right) \Big/ \sqrt{np(1-p)}$$
$$z_H = \left(x_0 - .5 - np\right) \Big/ \sqrt{np(1-p)}$$

وبالتالي

$$P(X = x_0) \approx P(x_0 - 0.5 \le X \le x_0 + 0.5)$$
$$= F(z_H) - F(z_L)$$

١٤-٥ تقريب بواسون

في حال لم يُطبّق التقريب الطبيعي، يتم استخدام تقريب بواسون إن أمكن عندما تكون n كبيرة و p صغيرة، وتصبح معلمة بواسون التي تمثل متوسط المتغير X على النحو التالي:

$$\theta = np$$

باستخدام توزيع بواسون، يصبح احتمال x:

$$P(x) = e^{-\theta}\, \theta^x \Big/ x! \qquad x = 0,1,\ldots$$

المثال (١٤-٢). ينبغي إنتاج الكثير من وحدات $n = 10$ في أحد المرافق التي تنتج 10% من الوحدات المعيبة. مع المتغير X، يكون عدد الوحدات المعيبة ومتوسط x وانحرافه المعياري كما هو مبين أدناه:

$$\begin{aligned}\mu &= np\\ &= 10 \times 0.1\\ &= 1.0\\ \sigma &= \sqrt{np(1-p)}\\ &= \sqrt{10(0.1)(0.9)}\\ &= 0.95\end{aligned}$$

ويكون التوزيع الاحتمالي للمتغير X على النحو التالي:

$$P(x) = \binom{10}{x} 0.1^x \, 0.9^{10-x}, \qquad x = 0,1,\ldots,10$$

ومن ثم يتم حساب احتمال الحصول على اثنتين من الوحدات المعيبة أو أكثر كما هو مبين أدناه:

$$\begin{aligned} P(X \ge 2) &= 1 - P(X \le 1) \\ &= 1 - \left[P(0) + P(1)\right] \end{aligned}$$

حيث إن

$$\begin{aligned} P(0) &= \binom{10}{0} 0.1^0 \, 0.9^{10} = 0.349 \\ P(1) &= \binom{10}{1} 0.1^1 \, 0.9^9 = 0.387 \end{aligned}$$

وبالتالي

$$\begin{aligned} P(X \ge 2) &= 1 - \left[0.349 + 0.387\right] \\ &= 0.264 \end{aligned}$$

المثال (١٤-٣). يُجري كيميائي مختبر تجربة $n = 8$ مرات وعدد النجاحات هو $x = 6$. ويصبح احتمال تقدير النجاح:

$$\hat{p} = x/n = 6/8 = 0.75$$

ويكون الانحراف المعياري s_p لتقدير الاحتمال p على النحو التالي:

$$\begin{aligned} s_p &= \left(\hat{p}\left(1 - \hat{p}\right)/n\right)^{0.5} \\ &= \sqrt{0.75(0.25)/8} \\ &= 0.153 \end{aligned}$$

المثال (١٤-٤). يتم تنفيذ الكمية المعدّة للإنتاج $n = 12$ وحدة كل يوم لمدة خمسة أيام متتالية مع وجود الوحدات المعيبة x_i في كل دفعة، وتكون دفعات النتائج الخمس على النحو التالي: (0، 1، 0، 2، 1). يتم حساب تقدير احتمال الوحدة المعيبة على النحو المبين أدناه:

$$\begin{aligned}\hat{p} &= \sum x_i \big/ (nm) \\ &= 4\big/(5\times 12) = 0.067\end{aligned}$$

ويكون الانحراف المعياري:

$$\begin{aligned}s_p &= \sqrt{0.067(0.933)\big/60} \\ &= 0.032\end{aligned}$$

المثال (١٤-٥). يسلّم أحد المورّدين 200 وحدة كل أسبوع لأحد المصانع ومن معلومات سابقة نعلم أن نسبة الوحدات المعيبة هي 0.05. يحتاج المصنع إلى 185 وحدة غير معيبة كل أسبوع ويريد تحديد احتمال هذا الحدث. إن احتمال الحصول على 185 وحدة غير معيبة أو أكثر هو مثل احتمال الحصول على 15 وحدة معيبة أو أقل.

باعتبار أن

$$p = 0.05 < 0.50$$

و

$$np = 200 \times 0.05 = 10 > 5$$

يمكن استخدام التقريب الطبيعي لتقريب الاحتمال، ويكون المتوسط والانحراف المعياري للمتغير X والذي يمثل عدد الوحدات المعيبة، على النحو المبين أدناه:

$$\begin{aligned}\mu &= np = 200 \times 0.05 = 10 \\ \sigma &= \sqrt{200 \times 0.05(0.95)} = 3.08\end{aligned}$$

كما يتم تقدير احتمال أن يكون المتغير X أقل أو يساوي 15 على النحو المبين أدناه:

$$\begin{aligned}P(X \le 15) &= P\Big(Z \le \Big[\big(15.5 - 10.0\big)\big/3.08\Big]\Big) \\ &= P(Z \le 1.79)\end{aligned}$$

عند تطبيق الجدول (٨-١) المبين في الفصل الثامن، يصبح الاحتمال التراكمي على النحو التالي:

$$\begin{aligned} P(X \leq 15) &\approx F(1.79) \\ &= 0.963 \end{aligned}$$

وبالتالي

P(185 وحدة غير معيبة أو أكثر) $\approx P(X \leq 15) = 0.963$

المثال (١٤-٦). عند إنتاج دليل هواتف، وبافتراض أن تقدير خطأ الصفحة هو 0,004. ما هو احتمال وجود خطأين أو أكثر في دليل مؤلف من 1000 صفحة؟ و يلاحظ أنه لا يُطبّق التقريب الطبيعي باعتبار أن $p = 0.004$ أقل من 0.5 و

$$np = 1000 \times 0.004 = 4 < 5$$

وبدلاً من ذلك يمكن استخدام تقريب بواسون باعتبار أن n كبيرة وp صغيرة. في هذه الحالة تصبح معلمة بواسون:

$$\theta = np = 1000 \times 0.004 = 4.0$$

ويتم حساب الاحتمال المطلوب على النحو أدناه:

$$\begin{aligned} P(X \geq 2) &= 1 - P(X \leq 1) \\ &= 1 - \left[P(0) + P(1)\right] \\ &= 1 - \left[e^{-4}4^0/0! + e^{-4}4^1/1!\right] \\ &= 1 - \left[0.0183 + 0.0733\right] \\ &= 0.908 \end{aligned}$$

١٤-٦ بيانات العيّنة

عند إجراء إحدى التجارب وتكرارها عدة مرات وتسجيل عدد النجاحات، تكون بيانات العيّنة على النحو التالي:

$n =$ عدد المحاولات

$x =$ عدد النجاحات

١٤-٧ تقدير المعلمة عند توفر بيانات العيّنة

باستخدام بيانات عيّنة حجمها n ونجاحات x، تكون طريقة مقدّر الإمكان الأعظم للاحتمال p على النحو التالي:

$$\hat{p} = x/n$$

ويصبح الانحراف المعياري للاحتمال p على النحو التالي:

$$s_p = \sqrt{\hat{p}(1-\hat{p})/n}$$

وعند إجراء التجربة ذات المحاولة n عدد m مرة، ويكون عدد النجاحات هو $(x_1, x_2, \ldots, x_m)$، يصبح تقدير الاحتمال p على النحو التالي

$$\hat{p} = \sum x_i/(nm)$$

وفي حال كانت التجربة $i-th$ ذات المحاولات n_i والنجاحات x_i، يصبح تقدير الاحتمال p على النحو التالي:

$$\hat{p} = \sum x_i / \sum n_i$$

١٤-٨ تقدير المعالم عند عدم توفر بيانات

على افتراض أن هناك حاجة لتطبيق توزيع ثنائي الحدين بحيث يكون عدد المحاولات n معلوماً، إلا أنَّ احتمال نجاح كل محاولة p مجهول. وعلى افتراض عدم توفر بيانات لتقدير الاحتمال p. في هذه الحالة قد يطلب المحلل استشارة فيما يتعلق بالقيمة المرجحة للمتغير x في المحاولات n. يُشار إلى هذا التقدير بالرمز $\tilde{x}$. عند استخدام المعادلة لقياس المنوال المحدد سابقاً، يتم تشكيل العلاقة التالية:

$$\tilde{x} = p(n+1)$$

وبالتالي يصبح تقدير احتمال p كما يلي:

$$\hat{p} = \tilde{x}/(n+1)$$

المثال (١٤-٧). على افتراض سيناريو أنه سيتم اختبار $n = 10$ محاولات واحتمال النجاح p مجهول. إضافة إلى ذلك، لا تتوفر بيانات العينة إلا أنَّ تقريب العدد المرجّح للنجاحات في عشر محاولات هو 3، وبالتالي $\tilde{x} = 3$، ويصبح تقدير احتمال كل محاولة :

$$\hat{p} = 3\big/(10+1) = 0.273$$

١٤-٩ ملخص

يُطبّق التوزيع ثنائي الحدين عندما تحدث محاولات n لإحدى التجارب ويبقى احتمال نجاح نتيجة ما في كل محاولة ثابتاً في جميع المحاولات. إنَّ المتغير العشوائي X هو عدد النجاحات في المحاولات n. وقد تم توضيح احتمال نجاحات X في المحاولات n، إلى جانب الاحتمال التراكمي للمتغير X أو عدد أقل من النجاحات. كما تم إدراج متوسط عدد النجاحات، وتباينه، وانحرافه المعياري، ومنواله. كما يمكن التوصل إلى تقدير احتمال نجاح كل محاولة عند توفر بيانات العينة وعند عدم توفرها.

الفصل الخامس عشر

التوزيع الهندسي
GEOMETRIC DISTRIBUTION

١٥-١ مقدمة

يُطبّق التوزيع الهندسي عند إجراء تجربة بصورة متكررة حتى الحصول على أول حالة نجاح. ويكون احتمال النجاح ذاته بالنسبة لجميع المحاولات. يمكن تعريف المتغير العشوائي على أنه عدد الإخفاقات إلى أن يتحقق النجاح الأول، ويتضمن مداه الأعداد الصحيحة من صفر وما وفوق. كما يمكن تصنيف المتغير العشوائي على أنه عدد المحاولات إلى أن يتحقق النجاح الأول ويتضمن مداه الأعداد الصحيحة من واحد وما فوق. تم توضيح كلا السيناريوهين في الفصل الحالي. وتحدث هذه الحالة على سبيل المثال، عندما تنتج إحدى العمليات الوحدات اللازمة لتلبية المعايير الهندسية المقبولة، ويتم تكرار العملية إلى أن يتم إنتاج وحدة مقبولة. وعندما يكون احتمال النتيجة الناجحة مجهولاً، يمكن استخدام بيانات العينة لتقدير هذا الاحتمال. في بعض الأحيان لا تتوفر بيانات العينة، ويقدم أحد الخبراء تقريباً عن التوزيع ويتم استخدام هذه البيانات لتقدير احتمال نجاح النتيجة. كما يبيّن هذا الفصل كيف أن التوزيع الهندسي هو التوزيع المنفصل الوحيد الذي يتميّز بخاصية فقدان الذاكرة.

١٥-٢ أساسيات

يحدث التوزيع الهندسي عند تكرار إحدى العمليات إلى أن نحصل على حالة نجاح، ويُعرّف المتغير العشوائي على أنه عدد الإخفاقات إلى أن يتحقق النجاح. وفي بعض الأحيان، يُعرّف المتغير العشوائي على أنه عدد المحاولات إلى أن يتحقق النجاح. يبيّن هذا الفصل كلا السيناريوهين.

١٥-٣ عدد الإخفاقات

تمثّل المعلمة p احتمال نجاح كل محاولة. كما يُصنّف المتغير العشوائي X والذي يمثل عدد الإخفاقات إلى أن يتحقق النجاح الأول كما هو مبيّن أدناه:

$$x = 0, 1, 2, \ldots$$

إنَّ الدالة التي تحدد احتمال المتغير X تكون على النحو التالي:

$$P(x) = p(1-p)^x, \qquad x = 0, 1, 2, \ldots$$

ويكون متوسط و تباين X على النحو المبين أدناه:

$$\mu_x = (1-p)\big/p$$
$$\sigma_x^2 = (1-p)\big/p^2$$

ومن ثم يتم التوصل إلى التوزيع التراكمي للمتغير X أو أقل على النحو التالي:

$$F(x) = 1-(1-p)^{x+1}, \qquad x = 0, 1, 2, \ldots$$

١٥-٤ بيانات العيّنة

غالباً ما تُستخدم بيانات العيّنة لتقدير المعلمة p. وعندما يكون المتغير العشوائي $X =$ عدد الإخفاقات إلى أن يتحقق النجاح، يتم إجراء التجربة m مرة وتُصنّف النتائج على شكل $(x_1, x_2, \ldots, x_m)$.

١٥-٥ تقدير المعلمة عند توفر بيانات العيّنة

تُستخدم بيانات العيّنة أعلاه لتقدير احتمال نجاح كل محاولة على النحو التالي:

أولاً: يتم حساب متوسط قيمة x لعدد m من المرات كما هو مبين أدناه:

$$\bar{x} = \sum x_i \big/ m$$

ثانياً: يتم التوصل إلى تقدير الاحتمال p على النحو التالي:

$$\hat{p} = 1\big/(\overline{x} + 1)$$

المثال (١٥-١). لدى فريق التنقيب عن النفط احتمال النجاح $p = 0.2$ لكل عملية حفر. باستخدام المتغير X على أنه عدد الإخفاقات حتى الحصول على نجاح وبالتالي يكون احتمال المتغير X على النحو التالي:

$$P(x) = 0.2(0.8^{x}), \qquad x = 0,1,\ldots$$

وفي هذه الحالة:

$$\begin{aligned} P(0) &= 0.200 \\ P(1) &= 0.160 \\ P(2) &= 0.128 \end{aligned}$$

وما إلى ذلك. ويكون احتمال ضرورة حدوث إخفاقين أو أقل لتحقيق النجاح على النحو التالي:

$$F(2) = P(X \le 2) = 0.200 + 0.160 + 0.128 = 0.488$$

ومن ثم يتم حساب متوسط المتغير X، وتباينه، وانحرافه المعياري على النحو المبين أدناه:

$$\begin{aligned} \mu_{x} &= (1 - 0.2)\big/0.2 = 4.00 \\ \sigma_{x}^{2} &= (1 - 0.2)\big/0.2^{2} = 20.00 \\ \sigma_{x} &= \sqrt{20.00} = 4.47 \end{aligned}$$

المثال (١٥-٢). لنفترض أن لدينا 5 عيّنات $m = 5$ من البيانات الهندسية ونتج منها القيم التالية للمتغير X والذي يمثل عدد الإخفاقات إلى أن يتحقق النجاح كما يلي: [3، 7، 5، 4، 5]، عندئذ يكون متوسط العينة: $\overline{x} = 4.80$، ومن ثم يصبح تقدير p على النحو المبين أدناه:

$$\hat{p} = 1\big/(4.80 + 1) = 0.172$$

١٥-٦ عدد المحاولات

تمثّل معلمة التوزيع الهندسي p احتمال نجاح كل محاولة بحيث تبقى ثابتة مع كل المحاولات. وعندما يكون عدد المحاولات n اللازم لتحقيق النجاح الأول هو المتغير العشوائي، يكون مدى n على النحو التالي:

$$n = 1, 2, \ldots$$

و يكون احتمال n على النحو المبين أدناه:

$$P(n) = p(1-p)^{n-1}, \qquad n = 1, 2, \ldots$$

ومن ثم يكون الاحتمال التراكمي لـ n أو أقل على النحو التالي:

$$F(n) = 1-(1-p)^{n}, \qquad n = 1, 2, \ldots$$

ويكون متوسط وتباين n على النحو المبين أدناه:

$$\mu_x = 1/p$$
$$\sigma_x^2 = (1-p)/p^2$$

باعتبار أن $n = x + 1$

$$\mu_n = \mu_x + 1$$
$$\sigma_n^2 = \sigma_x^2$$

المثال (١٥-٣). بافتراض أن العملية التي يكون فيها احتمال النجاح $p = 0.7$، يريد المحلل إيجاد احتمال المحاولات n إلى أن يتحقق النجاح. فتصبح دالة الاحتمال كما يلي:

$$P(n) = 0.7(1-0.7)^{n-1}, \qquad n = 1, 2, \ldots$$

حيث إن

$$P(1) = 0.700$$
$$P(2) = 0.210$$
$$P(3) = 0.063$$

وهكذا.

ويكون الاحتمال التراكمي بأن تكون $n = 3$ أو أقل على النحو التالي:

$$\begin{aligned} F(3) &= 1 - (1 - 0.7)^3 \\ &= 1 - 0.027 \\ &= 0.973 \end{aligned}$$

ويكون متوسط n وتباينها وانحرافها المعياري على النحو المبيّن أدناه:

$$\begin{aligned} \mu_n &= 1/0.7 = 1.428 \\ \sigma_n^2 &= (1 - 0.7)/0.7^2 = 0.612 \\ \sigma_n &= \sqrt{0.612} = 0.782 \end{aligned}$$

١٥-٧ بيانات العيّنة

غالباً ما تُستخدم بيانات العيّنة لتقدير المعلمة p. عندما يكون المتغير العشوائي n (عدد المحاولات إلى أن يتحقق النجاح)، حيث يتم إجراء التجربة m مرة وتُصنّف النتائج على شكل $\left(x_1, x_2, \ldots, x_m\right)$.

١٥-٨ تقدير المعلمة عند توفر بيانات العيّنة

يتم استخدام بيانات العيّنة لتقدير احتمال نجاح كل محاولة على النحو التالي:

أولاً: يتم حساب قيمة متوسط n لعدد m من المرات على النحو المبين أدناه:

$$\bar{n} = \sum n_i / m$$

ثانياً: يتم التوصل إلى تقدير p على النحو التالي:

$$\hat{p} = \frac{1}{\overline{n}}$$

المثال (١٥-٤). على افتراض أنه قد تمَّ أخذ 8 عيّنات من n (عدد المحاولات إلى أن يتحقق النجاح) لتقدير احتمال نجاح كل محاولة مع النتائج (1، 3، 2، 1، 1، 2، 1، 2). وبالتالي فإن متوسط عدد المحاولات هو: $\overline{n} = 13/8 = 1.615$، ويكون تقدير p كما يلي:

$$\hat{p} = 1\big/1.625 = 0.615$$

١٥-٩ تقدير المعلمة عند عدم توفر بيانات العيّنة

في بعض الأحيان يريد المحلل تطبيق التوزيع الهندسي إلا أنه لا تتوفر لديه بيانات عن العينة لتقدير احتمال النجاح. في هذه الحالة يقدم أحد الخبراء تقريباً من النوع التالي: "إنَّ الاحتمال هو α حيث إن عدد المحاولات حتى النجاح هو $n\alpha$ وذلك بافتراض أنَّ:

$$P(n \le n\alpha) = \alpha$$

مع ملاحظة أنَّ

$$\alpha = F(n\alpha) = 1 - (1-p)^{n}\alpha$$

وباستخدام بعض العمليات الجبرية، يكون تقدير p على النحو التالي:

$$\hat{p} = 1 - (1-\alpha)^{1/n}\alpha$$

وبالتالي فإنه عند تقدير $n\alpha, \alpha$، نحصل على p.

المثال (١٥-٥). تريد إحدى المنشآت الإنتاجية تطبيق التوزيع الهندسي في أحد نماذج المحاكاة، وتحتاج إلى تقدير احتمال نجاح مُنتج جديد. لا توجد بيانات عن العيّنة لبناء التقدير على أساسها وتطلب الإدارة مشورة مهندس التصميم حيث يفيد المهندس بأنه ينبغي أن يكون هناك ثقة بنسبة 0.90% بحدوث النجاح في $n = 3$ محاولات أو أقل. وينتج عن هذا التقريب:

$$n\alpha = 3$$
$$\alpha = 0.90$$

وباستخدام العلاقة المبيّنة سابقاً، يصبح تقدير الاحتمال p على النحو التالي:

$$\hat{p} = 1 - (1 - 0.9)^{1/3} = 0.536$$

١٥-١٠ نسبة لكسيس

إنَّ نسبة لكسيس بالنسبة للمتغير X (عدد الإخفاقات إلى أن يتحقق النجاح) هي:

$$\tau = 1/p$$

وهي دائماً أكبر من واحد.

إلا أنَّ نسبة لكسيس بالنسبة للمتغير n (عدد المحاولات إلى أن يتحقق النجاح) هي:

$$\tau = (1 - p)/p$$

وهو أمرٌ غير مُقنع باعتبار أنَّها قد تكون أكثر أو أقل من واحد.

١٥-١١ خاصية فقدان الذاكرة

يتمتع التوزيع الهندسي بخاصية فقدان الذاكرة من النوع التالي حيث يبقى الاحتمال التراكمي لعدد المحاولات $n = n_0$ أو أقل ثابتاً بغض النظر عن نقطة بدء الحساب. نبيّن فيما يلي هذه العلاقة. ونلاحظ أولاً أنَّ الاحتمال المكمّل لعدد المحاولات n هو أكبر من n_0

$$\begin{aligned} P(n > n_0) &= 1 - P(n \le n_0) \\ &= 1 - 1 + (1 - P)^{n_0} \\ &= (1 - P)^{n_0} \end{aligned}$$

كما نلاحظ أنَّ الاحتمال المكمّل لعدد المحاولات n أكبر من $n_1 + n_0$، مع العلم أنه قد حدثت المحاولات n_1 عندما بدأ عدّ المحاولات:

$$\begin{aligned} P\left(n > n_1 + n_0 \mid n_1\right) &= \left(1 - p\right)^{n_0 + n_1} \Big/ \left(1 - p\right)^{n_1} \\ &= \left(1 - p\right)^{n_0} \end{aligned}$$

باعتبار أنَّ الاحتمالات المكمّلة هي ذاتها، فإنَّ التوزيع يتمتع بخاصية فقدان الذاكرة.

١٥-١٢ ملخص

يمثل المتغير العشوائي للتوزيع الهندسي عدد الإخفاقات إلى أن يتحقق النجاح، أو عدد المحاولات إلى أن يتحقق النجاح الأول بحيث يتم حساب احتمال كل منهما باستخدام المقاييس الإحصائية التي تتضمن المتوسط والانحراف المعياري. ويتم تقدير احتمال النجاح عندما تتوفر بيانات العينة، كما يتم تقديرها عند عدم توفر البيانات بحيث يتم تقريب هذا المتغير. كما أنَّ التوزيع الهندسي هو التوزيع المنفصل الوحيد الذي يتمتع بخاصية فقدان الذاكرة.

الفصل السادس عشر

توزيع باسكال
PASCAL DISTRIBUTION

١٦-١ مقدمة

يُعدُّ بليز باسكال (Blaise Pascal) وهو عالم رياضيات فرنسي بارز في القرن السابع عشر أول من صاغ توزيع باسكال. كما أنه غالباً ما يُشار إليه بالتوزيع ثنائي الحدين السالب. ويتم استخدامه عند إجراء تجربة قد تنجح أو تفشل مع الاحتمال p و $(1-p)$ على التوالي، ويسعى الباحث إلى تحقيق عدد k نجاح للتجربة، بحيث يكون المتغير العشوائي هو العدد الأدنى للإخفاقات التي تحدث لتحقيق k من النجاحات. عندئذ يُسمى هذا التوزيع بتوزيع باسكال. يهتم بعض المحللين الذين يستخدمون توزيع باسكال بتحديد المتغير العشوائي عندما يكون العدد الأدنى لمحاولات تحقيق النجاحات k. على سبيل المثال، عندما تحتاج إحدى المنشآت إلى إنتاج k من الوحدات الناجحة لتلبية طلب أحد العملاء ويكون احتمال نجاح الوحدة أقل من واحد، يعرف المتغير العشوائي على أنه عدد الإخفاقات إلى أن يتم إنجاز k من الوحدات الناجحة. يبيّن الفصل كيفية قياس احتمالات كل حالة. وعندما يكون احتمال نجاح كل محاولة مجهولاً، يتم استخدام بيانات العيّنة لتقدير الاحتمال. في بعض الأحيان لا تتوفر بيانات عن العينة، ويتم استخدام تقريب للتوزيع لتقدير الاحتمال.

١٦-٢ أساسيات

يُشار أحياناً إلى المتغير العشوائي لتوزيع باسكال بأنه عدد الإخفاقات إلى أن يتم تحقيق k من النجاحات، ويُشار إليه أحياناً بأنه عدد المحاولات إلى أن يتم تحقيق k من النجاحات. نبيّن فيما يلي كيفية حساب القياسات الاحتمالية لكل حالة.

١٦-٣ عدد الإخفاقات

عندما يكون المتغير العشوائي يمثل عدد الإخفاقات إلى أن يتم تحقيق k من النجاحات، يكون المتغير المُشار إليه بأنه X على النحو التالي:

$$x = 0,1,2,\ldots$$

يتم التوصل إلى احتمال X، واحتماله التراكمي كما هو مبيّن أدناه:

$$P(x) = \binom{k+x-1}{x} p^k \left(1-p\right)^x \qquad x = 0,1....$$
$$F(x) = \sum_{j=0}^{x} \binom{k+j-1}{j} p^k \left(1-p\right)^j \qquad x = 0,1....$$

إنَّ متوسط وتباين X يعرفان على النحو التالي:

$$\mu_x = k\ (1-p)/p$$
$$\sigma_x^2 = k\ (1-p)/p^2$$

كما يتم اشتقاق منوال X، المشار إليه بالرمز $\tilde{\mu}$، على النحو التالي:

ليكن $w = \left[k\,(1-p)-1\right]/p$

إذا كان w عدداً صحيحاً فإن $\tilde{\mu} = w$ و $1+w$

وفيما عدا ذلك $\tilde{\mu} =$ الحد الأدنى لـ $(1+w)$

١٦-٤ تقدير المعلمة عند توفر بيانات العيّنة

عندما تكون المعلمة p مجهولة، يمكن استخدام بيانات العينة لتقدير القيمة. يتم الحصول على البيانات عندما يتم جمع مشاهدات $\left(x_1, x_2, \ldots, x_m\right)$ بعدد m للمتغير X، وهو ما يمثل عدد الإخفاقات من كل عينة من عيّنات m. ثم يتم حساب متوسط قيمة x من العينة كما هو مبيّن أدناه:

$$\bar{x} = \sum x_i / m$$

ويكون تقدير الإمكان الأعظم للمعلمة p على النحو التالي:

$$\hat{p} = k/(\bar{x} + k)$$

١٦-٥ تقدير المعلمة عند عدم توفر البيانات

في بعض الأحيان يحتاج المحلل إلى تطبيق توزيع باسكال ولا تتوفر لديه بيانات عن العيّنة لتقدير p (احتمال النجاح). في هذه الحالة قد يطلب الباحث مشورة أحد الخبراء فيما يتعلق بالنموذج ويطلب تقريباً للقيمة المرجّحة للمتغير X. إنَّ هذا التقريب هو مثل منوال المتغير X والمشار إليه بالرمز $\tilde{x}$. باستخدام تقدير المنوال ومعرفة قيمة k مُسبقاً، يتم التوصل إلى تقدير p كما هو مبيّن أدناه. مع ملاحظة كيفية التوصل إلى المنوال فيما سبق باستخدام القيمة المؤقتة للمقدار w والمبيّنة أدناه:

$$w = \left[k(1-p)-1\right]/p$$

وعند استبدال المنوال بالقيمة w أي أن $\tilde{x} = w$، وتطبيق بعض العمليات الجبرية، ينتج التقدير التالي للاحتمال p:

$$\hat{p} = \left(k-1\right)/\left(k+\tilde{x}\right)$$

المثال (١٦-١). يسعى أحد الكيميائيين إلى إجراء خليط من المواد الكيميائية والحرارة، والذي يعطي نتيجة إيجابية مع احتمال $p = 0.5$ لكل محاولة. ويحتاج الكيميائي إلى مكائن للخلاطة $k = 3$ لإنجاز التجربة. يتم حساب المتغير X عندما يكون مساوياً لعدد المكائن التي تفشل إلى أن يتم إنجاز ثلاثة مركبات إيجابية، على النحو المبين أدناه:

$$P(x) = \binom{3+x-1}{x} 0.5^3 (1-0.5)^x \qquad x = 0,1,\ldots$$

ويتم التوصل إلى احتمال $x = 0$ أو 1 على النحو التالي:

$$P(0) = \binom{3+0-1}{0} 0.5^3 (1-0.5)^0 = 0.125$$

$$P(1) = \binom{3+1-1}{1} 0.5^3 (1-0.5)^1 = 0.187$$

واحتمال أن يكون X أقل من أو يساوي واحداً هو:

$$F(1) = 0.125 + 0.187 = 0.312$$

ومن ثم يكون المتوسط، و التباين، والانحراف المعياري على النحو التالي:

$$\mu_x = 3(1-0.5)/0.5 = 3.00$$
$$\sigma_x^2 = 3(1-0.5)/0.5^2 = 6.00$$
$$\sigma_x = \sqrt{6.00} = 2.45$$

المثال (١٦-٢). يتطلب أحد المجمعات وحدتين وبناء محل تجاري حسب الحاجة. إنَّ الوحدتين حساستان للغاية ولا ينتج التشغيل الآلي دائماً وحدات مقبولة، واحتمال قبول الوحدة مجهول. يحتاج المحل إلى تقدير إنتاج الوحدة المقبولة لتسعير طلب العمل على نحوٍ ملائم. خلال الأشهر الماضية، يستعيد المحل $m = 8$ طلبات مماثلة ويجمع عدد الإخفاقات التي حدثت. وقد تم توضيح ذلك على النحو التالي:

$$x_i = (2, 5, 3, 4, 6, 7, 2, 5)$$

يكون متوسط العينة بالنسبة لعدد الإخفاقات على النحو المبين أدناه:

$$\bar{x} = 34/8 = 4.25$$

و باستخدام المتوسط، يتم التوصل إلى احتمال إنتاج وحدة مقبولة على النحو التالي:

$$\hat{p} = 2/(4.25 + 2) = 0.32$$

المثال (١٦-٣). يقوم أحد الباحثين بإنشاء نموذج محاكاة ويحتاج إلى استخدام توزيع باسكال حيث إن العدد المطلوب من الوحدات المقبولة هو $k = 4$، واحتمال p الحصول على وحدة مقبولة بالنسبة لكل محاولة، مجهول. يطلب الباحث رأي العديد من زملاء العمل فيما يتعلق بالقيمة المرجّحة لعدد الوحدات غير المقبولة بالنسبة لهذه العملية. إنَّ التقريب النموذجي هو $\tilde{x} = 3$. ومن خلال هذا التقريب، يتم التوصل إلى تقدير الاحتمال p كما هو مبين أدناه:

$$\hat{p} = (4-1)/(4+3) = 0.428$$

١٦-٦ عدد المحاولات

بافتراض توزيع باسكال عندما يكون المتغير العشوائي عبارة عن عدد المحاولات إلى أن تتحقق k من النجاحات. في هذه الحالة، يُشار إلى المتغير بأنه n حيث إن:

$$n = k, k+1, \ldots$$

ويتم حساب احتمال n واحتماله التراكمي كما هو مبين أدناه:

$$P(n) = \binom{n-1}{k-1} p^k \left(1-p\right)^{n-k} \qquad n = k, k+1, \ldots$$

$$F(n) = \sum_{y=k}^{n} p(y). \qquad n = k, k+1, \ldots$$

ومن ثم يكون متوسط وتباين n كما يلي.

$$\mu_n = k/p$$
$$\sigma_n^2 = k(1-p)/p^2$$

باعتبار أن $n = (x+k)$، يرتبط متوسط n وانحرافه المعياري بما يقابلها عندما يكون x عدد الإخفاقات، كما هو مبين أدناه:

$$\mu_n = \mu_x + k$$

$$\sigma_n^2 = \sigma_x^2$$

١٦-٧ نسبة لكسيس

إنَّ نسبة لكسيس بالنسبة للمتغير X (عدد الإخفاقات إلى أن تتحقق k من النجاحات) هي:

$$\tau = 1/p$$

وبالتالي فهي دائماً أكبر من واحد. وإنَّ ذات المقياس بالنسبة للمتغير n (عدد المحاولات إلى أن تتحقق k من النجاحات) هو:

$$\tau = (1-p)/p$$

ويكون أحياناً أقل من واحد وأحياناً أكثر ولذلك فهو غير مفيد للاستخدام.

١٦-٨ تقدير المعلمة عند توفر بيانات العيّنة

عندما تكون المعلمة p مجهولة، يمكن تقديرها باستخدام بيانات العينة. على افتراض أنه قد تم ملاحظة العيّنات m عن المتغير n بشكلٍ عشوائي، وتحقيق النتائج $(n_1,\ldots,n_m)$. تعطي البيانات متوسط العينة n، وكذلك تقدير المعلمة p، كما هو موضح أدناه:

$$\bar{n} = \sum n_i / m$$
$$\hat{p} = k/\bar{n}$$

المثال (١٦-٤). يريد أربعون بالمائة من الناخبين إجراء استفتاء، ويبحث أحد المراسلين عن خمسة منهم لكتابة مقال مميّز. نبيّن فيما يلي عدد الناخبين الذي يجب على المراسل مقابلتهم. في هذه الحالة يمثل n عدد الناخبين الذين قام المراسل بإجراء مقابلات معهم (عدد المحاولات) و $p = 0.40$ هو احتمال إجراء الناخب للاستفتاء. يكون احتمال n على النحو التالي:

$$P(n) = \binom{n-1}{5-1} 0.4^5 (0.6)^{n-5}. \qquad n = 5,6,\ldots$$

نلاحظ أنَّ

$$P(5) = \binom{5-1}{5-1} 0.4^5 (0.6)^0 = 0.0102$$
$$P(6) = \binom{6-1}{5-1} 0.4^5 (0.6)^1 = 0.0307$$

وهكذا، يصبح احتمال أن يكون $n = 6$ أو أقل على النحو التالي:

$$P(n \le 6) = 0.0102 + 0.0307 = 0.0409$$

ومن ثم فإنَّ متوسط وتباين n يعرفان على النحو المبيّن أدناه:

$$\mu_n = 5/0.4 = 12.50$$

$$\sigma_n^2 = 5\left(1-0.4\right)\big/0.4^2 = 5.56$$

$$\sigma_n = \sqrt{5.56} = 2.36$$

كما نلاحظ أنَّ متوسط وتباين X والذي يمثل عدد الإخفاقات معرفان على النحو التالي:

$$\mu_x = (12.5-5) = 7.5$$
$$\sigma_x^2 = 5.56$$

المثال (١٦-٥). على افتراض أنه قد أُقيمت مسابقة داخلية حيث يتنافس $m = 3$ لاعبين ليحقّقوا أقل عدد من محاولات ضرب خمس كرات فوق السياج $(k = 5)$ وعلى افتراض أنَّ جميعهم متساوون. إنَّ عدد المحاولات بالنسبة للمشاركين الثلاثة يكون على النحو التالي: [15, 12, 14]. يستفسر أحد المراسلين عن احتمال أن يضربوا الكرة فوق السياج، باعتبار أن متوسط عدد المحاولات هو $\bar{n} = 13.67$، يصبح تقدير الاحتمال $\hat{p} = 5/13.67 = 0.365$.

١٦-٩ ملخص

إنَّ توزيع باسكال هو التوزيع الذي يكون فيه المتغير العشوائي يمثل عدد الإخفاقات إلى أن تتحقق k من النجاحات لحدثٍ ما، حيث إن احتمال النجاح ثابت بالنسبة لجميع المحاولات. وقد يكون المتغير العشوائي يمثل عدد المحاولات إلى أن تتحقق k من النجاحات. يتم حساب احتمال عدد الإخفاقات (أو المحاولات) بسهولة، وكذلك المتوسط والانحراف المعياري. وعندما يكون احتمال النجاح مجهولاً، يتم استخدام بيانات العيّنة لتقدير الاحتمال. وعند عدم توفر البيانات، يساعد تقدير العدد المرجّح للمحاولات في تقدير احتمال نجاح كل محاولة.

الفصل السابع عشر

توزيع بواسون
POISSON DISTRIBUTION

١٧–١ مقدمة

تمَّ تسمية توزيع بواسون نسبة إلى العالم سيميون بواسون (Simmeon Poisson)، وهو عالم رياضيات فرنسي بارز خلال أوائل القرن التاسع عشر. يُطبّق هذا التوزيع عندما يكون المتغير العشوائي منفصلاً ويمثّل عدد الأحداث التي تقع في وحدة المقياس، مثل وحدة الزمن أو وحدة المساحة. إنَّ معدل وقوع الأحداث ثابت وإنَّ عدد الأحداث يمثل أعداداً صحيحة من الصفر وما فوق. يُستخدم هذا التوزيع كثيراً في دراسة نُظم صفوف الانتظار، وكذلك في ضبط العملية الإحصائية. ففي نظام صفوف الانتظار، يتمثل المتغير العشوائي في عدد وحدات الوصول إلى أحد الأنظمة في وحدة الزمن، وكذلك في عدد وحدات المغادرة في وحدة الزمن بالنسبة لأحد المرافق الخدمية المشغولة بصورةٍ مستمرة. وخلال ضبط العملية الإحصائية، يكون المتغير العشوائي هو عدد العيوب الموجودة في وحدة مُنتج. غالباً ما يُوصف التوزيع بأنه وحدة مقياس واحدة، إلا أنه يمتد بسهولة إلى عدة وحدات قياس. يُستخدم هذا التوزيع في كثيرٍ من الأحيان بدلاً من التوزيع الطبيعي وذلك عندما يكون العدد المتوقع للأحداث منخفضاً نسبياً.

١٧–٢ أساسيات

يحتوي توزيع بواسون على المعلمة θ، وتمثل قياس عدد الأحداث التي تقع في وحدة مقياس. قد يكون المقياس زمناً، أو مساحة، أو حجماً، و هكذا. إنَّ المتغير العشوائي X يتضمن الأعداد الصحيحة صفراً أو أكبر. وقد تمَّ إدراج الدالة الاحتمالية للمتغير X أدناه:

$$P(x) = \theta^x e^{-\theta} \big/ x! \qquad x = 0,1,2,\ldots$$

ومنها فإنَّ متوسط و تباين X يكونان على النحو التالي:

$$\mu = \theta$$
$$\sigma^2 = \theta$$

كما يتم التوصل إلى منوال X ، المُشار إليه بالرمز $\tilde{\mu}$ ، على النحو التالي:

إذا كان θ عدداً صحيحاً: θ و $\tilde{\mu} = \theta - 1$

وفيما عدا ذلك الحد الأدنى للمعلمة $(\theta) = \tilde{\mu}$

١٧-٣ نسبة لكسيس

إنَّ نسبة لكسيس بالنسبة لتوزيع بواسون تكون على النحو المبين أدناه:

$$\tau = \theta \big/ \theta = 1$$

١٧-٤ تقدير المعلمة عند توفر بيانات العيّنة

عندما يتم جمع العيّنات n بالنسبة للمتغير X على شكل: $(x_1, \ldots, x_n)$ ويكون متوسط العينة $\overline{x}$ ، فإنَّ مقدر الإمكان الأعظم للمعلمة θ هو:

$$\hat{\theta} = \overline{x}$$

١٧-٥ تقدير المعلمة عند عدم توفر بيانات

عندما يكون هناك حاجة إلى تقدير قيمة المعلمة θ ولا تتوفر بيانات عن العينة، فقد تُستخدم العلاقة بالمنوال على النحو التالي. يطلب الباحث مشورة أحد الأشخاص المطلعين على النظام قيد الدراسة والذي يقدم بدوره تقريباً للقيمة المرجحة (المنوال) للمتغير X ، والمشار إليها بالرمز $\tilde{x}$. باعتبار أنَّ المنوال والمعلمة مرتبطان على الشكل التالي:

$$\theta \quad \text{و} \quad \tilde{\mu} = \theta - 1$$

فإنَّه ينتج عن استبدال $\tilde{x}$ بالمقدار $\tilde{\mu}$ تقدير المنوال على النحو المبين أدناه:

$$\hat{\theta} \approx \tilde{x} + 0.5$$

١٧-٦ الارتباط الأسي

يرتبط المتغير العشوائي المنفصل لتوزيع بواسون X بالمتغير العشوائي المتصل للتوزيع الأسي T على النحو التالي:

عندما تكون وحدة المقياس هي وحدة زمن، فإنَّ القيمة المتوقعة للمتغير X هي:

$E(x) = \mu = \theta$ = العدد المتوقع للأحداث في وحدة الزمن

يتم توزيع الزمن بين الأحداث والمُشار إليه بالمتغير t توزيعاً أسياً، وتكون الدالة الاحتمالية على النحو التالي:

$$f(t) = \theta e^{-\theta t}, \qquad t > 0$$

وتصبح القيمة المتوقعة للمتغير t:

$E(t) = 1/\theta$ =الزمن المتوقع بين الأحداث

المثال (١٧-١). يصل الزبائن إلى المطعم صباح يوم السبت بمعدل 5 زبائن كل 30 دقيقة، خلال هذه الفترة الزمنية يكون متوسط وتباين عدد الزبائن الوافدين على النحو التالي:

$$\mu = 5.0$$
$$\sigma^2 = 5.0$$

ومن ثم يتم حساب احتمال عدد مرات الوصول x خلال 30 دقيقة كما هو مبيّن أدناه:

$$P(x) = 5^x e^{-5}/x!, \qquad x = 0,1,2,\ldots$$

نلاحظ أن

$$P(0) = 5^0 e^{-5}/0! = 0.007$$
$$P(1) = 5^1 e^{-5}/1! = 0.034$$
$$P(2) = 5^2 e^{-5}/2! = 0.084$$

وعليه يصبح احتمال أن يكون المتغير X اثنين أو أقل على النحو التالي:

$$P(X \leq 2) = 0.007 + 0.034 + 0.084 = 0.125$$

المثال (١٧-٢). على افتراض أنَّ قيمة المعلمة في نظام بواسون مجهولة، ويحصل المحلل على العينات $n = 8$ للمتغير X كما يلي $x : \left[3,6,2,5,3,5,2,7\right]$ بناءً على بيانات العينة هذه يحسب المحلل المتوسط $\bar{x} = 4.125$ وبالتالي يصبح تقدير المعلمة $\hat{\theta} = 4.125$.

المثال (١٧-٣). يريد أحد الباحثين استخدام توزيع بواسون في نموذج محاكاة و ليس لديه تقدير معلمة بواسون ولا بيانات عن العينة أيضاً. يقدم أحد المطلعين على النظام تقريبات عن القيمة المرجحة للمتغير X على شكل $\tilde{x} = 3.0$. وباستخدام هذا التقريب، يصبح تقدير المعلمة:

$$\hat{\theta} = 3.0 + 0.5 = 3.5$$

المثال (١٧-٤). يصل العملاء إلى البنك في صباح السبت عبر توزيع بواسون بمعدل 6 عملاء بالساعة. يمكن التعبير عن معلمة بواسون على النحو التالي:

$\theta = 6$ عملاء بالساعة

أو

$\theta = 1$ عميل كل 10 دقائق

يكون الزمن بين عدد مرات وصول العملاء المشار إليه بالمتغير t أسي مع الزمن المتوقع على النحو التالي:

$$E(t) = 1/\theta$$

إذا كانت وحدة الزمن ساعة فإن

$E(t) = 1/6 = 0.167$ ساعة

وإذا كانت وحدة الزمن 10 دقائق فإن

$E(t) = 1/1 = 1.0$ دقيقة

باستخدام الساعات على أنها وحدة الزمن، تكون المعلمة $\theta = 6$ كل ساعة، وتكون دالة الكثافة الاحتمالية للمتغير t على النحو التالي:

$$f(t) = 6e^{-6t}, \qquad t > 0$$

١٧-٧ توزيع بواسون مع عدة وحدات

غالباً ما يُطبّق المتغير العشوائي لتوزيع بواسون مع وحدات متعددة المقياس n حيث إن n أكبر من الصفر. عندما يكون $n = 1$، تكون وحدة المقياس مثل المتغير العشوائي لتوزيع بواسون المبيّن أعلاه، وعند تحديد المعلمة θ من خلال وحدة مقياس واحدة، يُشار إلى المعلمة التي تُطبّق من خلال وحدات المقياس n ويتم الحصول عليها على النحو المبين أدناه:

$$\theta_n = \theta \times n$$

كما يبقى المتغير العشوائي مثل المتغير X حيث يكون المدى المقبول:

$$x = 0, 1, \ldots$$

كذلك تم إدراج الدالة الاحتمالية للمتغير X في وحدات المقياس n على النحو المبين أدناه:

$$P_n(x) = \theta_n\, e^{-\theta n} \big/ x!, \qquad x = 0, 1, 2, \ldots$$

ويكون متوسط وتباين المتغير X في وحدات المقياس n على النحو التالي:

$$\mu = \theta_n$$
$$\sigma^2 = \theta_n$$

المثال (١٧-٥). باعتبار المتغير العشوائي لتوزيع بواسون حيث تكون قيمة المعلمة لوحدة مقياس واحدة $\theta = 2.0$. نبيّن فيما يلي اختلاف بعض القياسات الإحصائية عندما يكون $n = 0.5, 1.0, 1.5$، على سبيل المثال.

$\mu = 1.0$ عندما يكون $n = 0.5$

$\mu = 2.0$ عندما يكون $n = 1.0$

$\mu = 3.0$ عندما يكون $n = 1.5$

$$\begin{array}{ll} P_{0.5}(x) = 1.0\,e^{-1.0x} \big/ x!, & x = 0,1,2,\ldots \\ P_{1.0}(x) = 2.0\,e^{-2.0x} \big/ x!, & x = 0,1,2,\ldots \\ P_{1.5}(x) = 3.0\,e^{-3.0x} \big/ x!, & x = 0,1,2,\ldots \\ P_{0.5}(0) = 0.368, & P_{0.5}(1) = 0.368 \\ P_{1.0}(0) = 0.135, & P_{1.0}(1) = 0.271 \\ P_{1.5}(0) = 0.050, & P_{1.5}(1) = 0.149 \end{array}$$

المثال (١٧-٦). في أحد معامل تصنيع الزجاج، يكون معدل متوسط الخطأ في لوح زجاج بقياس قدم مربّع 0,001. يتم التوصل إلى احتمال الخطأ في لوح زجاج بقياس 16 قدماً مربعاً على النحو التالي. بداية نلاحظ أن إعدادات المعلمة بالنسبة للوح زجاجي بقياس قدم مربع و16 قدماً مربعاً تكون على النحو المبين أدناه:

$\theta = 0.001$ بالنسبة للوح زجاجي بقياس قدم مربع

$\theta_{16} = 0.016$ بالنسبة للوح زجاجي بقياس 16 قدماً مربعاً

لذا يتم التوصل إلى احتمال وجود صفر خطأ في لوح زجاجي بقياس 16 قدماً مربعاً على النحو المبين أدناه:

$$P_{16}(0) = 0.016^0 e^{-0.016/0!} = 0.984$$

وبالتالي يصبح احتمال وجود خطأ واحد أو أكثر في لوح زجاجي بقياس 16 قدماً مربعاً على النحو التالي:

$$P_{16}(X \geq 1) = 1 - 0.984 = 0.016$$

١٧-٨ ملخص

يُطبّق توزيع بواسون عندما يكون عدد الأحداث في وحدة مقياس منخفضاً نسبياً. إنَّ المتغير العشوائي هو متغير منفصل ويتضمن جميع الأعداد الصحيحة من صفر وأكبر، ويرتبط العدد المنفصل للأحداث في وحدة زمن بالزمن المتصل الأسي بين الأحداث. غالباً ما يتم وصف التوزيع مع وحدة مقياس واحدة، إلا أنه يمتد بسهولة إلى عدة وحدات. يمكن تقدير معلمة توزيع بواسون باستخدام بيانات العينة. وعند عدم توفر بيانات العينة، يمكن تقدير قيمة المعلمة أيضاً.

الفصل الثامن عشر

التوزيع الهندسي الزائد (فوق الهندسي)
HYPER GEOMETRIC DISTRIBUTION

١٨-١ مقدمة

يُطبّق التوزيع الهندسي الزائد عندما يكون لدى مجتمع الدراسة الذي يبلغ حجمه N عناصر تحمل العلامة D، وقد نتج عن عيّنة العناصر n التي أُخذت دون استبدال، العناصر التي تحمل العلامة x. يختلف هذا التوزيع عن توزيع ثنائي الحدين حيث إن حجم المجتمع غير محدود وقد تم أخذ العينات من خلال الاستبدال. غالباً ما يُطبّق التوزيع الهندسي الزائد في تطبيقات الجودة عندما يُؤخذ الكثير من عناصر N ذات العيوب D (الكمية مجهولة) وكذلك أخذ عينة n دون استبدال. تساعد نتيجة العينة ذات العيوب x في ضبط جودة الكمية.

١٨-٢ أساسيات

تكون الدالة الاحتمالية للمتغير X على النحو التالي:

$$P(x) = \binom{N-D}{n-x}\binom{D}{x}\Big/\binom{N}{n} \qquad x = 0,\ldots,\min(n,D)$$

ويكون متوسط وتباين X على النحو المبيّن أدناه:

$$\mu = n\frac{D}{N}$$

$$\sigma^2 = n\left[D/N\right]\left[1-D/N\right]\left[N-n\right]/\left[N-1\right]$$

١٨-٣ تقدير المعلمة عند توفر بيانات العيّنة

يريد المحلل تطبيق التوزيع الهندسي الزائد في أحد التطبيقات بحيث يكون حجم الكمية والعينة معلومين، إلا أنَّ عدد العيوب في الكمية مجهول. خلال فترات m الماضية، وعلى اعتبار أن N و n ثابتان في الحجم، يكون عدد العيوب من العينة على النحو التالي: $(x_1,\ldots,x_m)$. ويصبح متوسط عدد عيوب كل فترة $\overline{x}$. باستخدام العلاقة بين D,N,x و n، يكون تقدير احتمال العيب على النحو التالي:

$$\frac{\hat{D}}{N} = \overline{x}/n$$

وبالتالي بالنسبة لحجم الكمية N، يصبح تقدير متوسط عدد العيوب على النحو التالي:

$$\overline{D} = N\overline{x}/n$$

١٨-٤ التقدير الثنائي

في حال كان حجم العينة n (عيّنات دون استبدال) صغيراً مقارنةً بالكمية N (حجم المجتمع)، يمكن استخدام توزيع ثنائي الحدين لتقدير احتمال التوزيع الهندسي الزائد وذلك بجعل $p = D/N$ مع حجم العينة n.

المثال (١٨-١). تصل شحنة من أحد الموردين إلى مصنع يضم الكثير من وحدات $N = 8$، وقد تم أخذ عينة من وحدتين $n = 2$ دون استبدال لإيجاد أي عيب. إذا كان عدد العيوب في الكمية هو $D = 1$، فإن احتمال إيجاد عيوب X يكون على النحو التالي:

$$P(x) = \binom{7}{2-x}\binom{1}{x}\Big/\binom{8}{2} \qquad x = 0,1$$

وتصبح الاحتمالات:

$$P(0) = \binom{7}{2-0}\binom{1}{0}\Big/\binom{8}{2} = 0.75$$
$$P(1) = \binom{7}{2-1}\binom{1}{1}\Big/\binom{8}{2} = 0.25$$

ويتم حساب متوسط وتباين X على النحو المبين أدناه:

$$\begin{aligned}\mu &= 2\times 1/8 = 0.25\\ \sigma^2 &= 2\times 1/8\left[1-1/8\right]\left[8-2\right]/\left[8-1\right] = 0.1875\end{aligned}$$

المثال (١٨-٢). على مدى ستة أسابيع، تصل شحنة إلى مصنع مع وحدات $N = 8$ وتم أخذ عينة $n = 2$ دون استبدال. ومن معلومات سابقة على مدى ستة أسابيع كان عدد العيوب كالتالي: (0، 1، 0، 0، 1، 0)، وهو ما يعطي متوسط العيوب $\overline{x} = 0.333$ لكل شحنة. لذا يكون تقدير العيوب بالنسبة لحجم الشحنة على النحو التالي:

$$\frac{\hat{D}}{N} = x/n = 0.333/2 = 0.167$$

وبالتالي فإن $\overline{D} = 0.167\times 8 = 1.33$ هو تقدير متوسط عدد العيوب في شحنة بحجم $N = 8$.

المثال (١٨-٣). يريد حارس الحديقة تقدير عدد الأسماك الكبيرة في البحيرة الصغيرة. اصطاد الحارس $D = 20$ سمكة ووضع لهم علامات. بعد وضع هذه العلامات، تم إعادة الأسماك ذات العلامات إلى البحيرة. وبعد فترةٍ قصيرة، اصطاد الحارس $N = 10$ سمكة أخرى ووجد $x = 2$ تحمل علامات عليه، يتم التوصل إلى تقدير عدد الأسماك الكبيرة N في البحيرة على النحو التالي:

$$\begin{aligned}E(x) &= nD/N\\ \hat{N} &= nD/x\\ &= 10\times 20/2 = 100\end{aligned}$$

المثال (١٨-٤). تم استلام شحنة بحجم $N = 10$ من أحد الموردين في المصنع وكان عدد العينات التي تم أخذها دون استبدال يساوي $n = 5$. على افتراض وجود العيوب $D = 8$ في الشحنة، يريد المحلل إيجاد احتمال العيوب x في العينة. تكون الدالة الاحتمالية للمتغير X في التوزيع الهندسي الزائد على النحو المبين أدناه:

$$P(x) = \binom{92}{5-x}\binom{8}{x}\Big/\binom{100}{5}$$

إنَّ الحسابات أعلاه صعبة بعض الشيء وذلك بسبب القيمة الكبيرة لـ $N = 100$. باعتبار أنَّ نسبة n على N صغيرة ($5/100 = 0.05$)، يمكن استخدام توزيع ثنائي الحدين لتقريب الاحتمالات المطلوبة. إنَّ الدالة الاحتمالية ثنائية الحدين للمتغير X تكون على الشكل التالي:

$$P(x) = \binom{5}{x} 0.08^x 0.92^{5-x}$$

نبيّن فيما يلي مقارنة بين احتمالات التوزيع الهندسي الزائد وتوزيع ثنائي الحدين:

توزيع ثنائي الحدين	التوزيع الهندسي الزائد	x
0.659	0.653	0
0.287	0.297	1
0.050	0.047	2
0.004	0.003	3
0.000	0.000	4
0.000	0.000	5

١٨–٥ ملخص

يُطبّق التوزيع الهندسي الزائد عندما تتضمن شحنة ذات حجم محدود بعض العناصر ذات خصائص معينة ويتم أخذ عينة صغيرة دون استبدال من الشحنة للحصول على عدد العناصر ذات الخاصية المعينة، وغالباً ما يُستخدم التوزيع في مجال تطبيقات ضمان الجودة.

الفصل التاسع عشر

التوزيع الطبيعي الثنائي
BIVARIATE NORMAL DISTRIBUTION

١٩-١ مقدمة

اهتم العديد من العلماء في وصف التوزيع الطبيعي الثنائي على مرّ السنين. في عام 1998، وصف مونتيرا جانتارافاريرات وثوموبولوس (Montira Jantaravareerat and N. Thomopoulos)، طريقة لتقدير الاحتمال التراكمي للتوزيع. وفي هذا الفصل يتم توضيح طريقة جديدة لحساب الاحتمال التراكمي المشترك. يحتوي التوزيع الطبيعي الثنائي على متغيرين X_1, X_2 المرتبطين بصورة مشتركة وله خمسة معالم هي $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$. يتم توزيع التوزيعات الهامشية توزيعاً طبيعياً، وعندما تكون قيمة أحد المتغيران معلومة، يتم توزيع المتغير الآخر توزيعاً طبيعياً. يتم تحويل المتغيران إلى مجموعة جديدة وهي Z_1, Z_2 المرتبطين بصورة مشتركة من خلال التوزيع الطبيعي القياسي الثنائي. من السهل تطبيق المتغيرين الأخيرين بشكلٍ رياضي في الحسابات. كما تمَّ وضع طريقة التقريب في هذا الفصل لحساب الاحتمال المشترك للمتغيرين. كما تمَّ إدراج القيم الجدولية وعرض أمثلة لتوضيح تطبيقات التوزيع.

١٩-٢ أساسيات

يبدأ الفصل بوصف التوزيع الطبيعي الثنائي والمتغيرين X_1, X_2 ومعالمه الخمسة. تُؤخذ المعالم من التوزيعات الهامشية للمتغيرين X_1, X_2، والتي يتم توزيعها توزيعاً طبيعياً. وعند تحديد قيمة لأحد المتغيرين، يكون للمتغير الآخر توزيع يتم توزيعه توزيعاً طبيعياً بمتوسط وانحراف معياري محددين. ونظراً لصعوبة الحسابات التي تستخدم المتغيرين X_1, X_2، يتم تحويل المتغيرين إلى نظيرهما Z_1, Z_2، المأخوذين من التوزيع الطبيعي القياسي الثنائي كونهما أسهل حسابياً للتطبيق. كما يتم وصف

التوزيعات الهامشية والشرطية للمتغيرين Z_1, Z_2 باعتبار أنه لا توجد صيغة محددة لحساب الاحتمال المشترك للزوج (k_1, k_2) اللذين هما قيمتان معيّنتان للمتغيرين (Z_1, Z_2)، كما تم وضع طريقة للتقريب في هذا الفصل. وباستخدام هذا التقريب، يتم إدراج بعض القيم الجدولية وعرض بعض الأمثلة لتوضيح كيفية استخدام الجداول.

١٩-٣ التوزيع الطبيعي الثنائي

عندما يتم توزيع المتغيرين X_1, X_2 توزيعاً طبيعياً ثنائياً، تحتوي علاقتهما المشتركة على خمسة معالم $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$، وتكون الدالة معرفة على النحو التالي، حيث يرتبط معلم المتوسط والانحراف المعياري بالتوزيع الهامشي للمتغيرين.

$$\left(X_1, X_2\right) \sim BVN(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$$

كذلك تم إدراج الدالة الاحتمالية المشتركة على النحو المبين أدناه:

$$f(x_1, x_2) = 1\Big/\Big[2\pi\sigma_1\sigma_2\sqrt{(1-p^2)}\exp\Big[-w\big/2(1-p^2)\Big]$$

$$w = \Big[\big(x_1 - \mu_1\big)\big/\sigma_1\Big]^2 + \Big[\big(x_2 - \mu_2\big)\big/\sigma_2\Big]^2 - 2\rho\Big[\big(x_1 - \mu_1\big)\big(x_2 - \mu_2\big)\big/\sigma_1\sigma_2\Big]$$

١٩-٤ التوزيعات الهامشية

فيما يلي التوزيعات الهامشية لكل متغير يتبع توزيعاً طبيعياً بحيث تكون على الشكل التالي:

$$X_1 \sim N(\mu_1, \sigma_1^2)$$
$$X_2 \sim N(\mu_2, \sigma_2^2)$$

نلاحظ أنَّ μ_1 و σ_1 هما المعلمتان الهامشيتان للمتغير X_1، في حين أن μ_2 و σ_2 هما المعلمتان الهامشيتان للمتغير X_2. والترابط بين المتغيرين هو ρ ويتم حسابه على النحو التالي:

$$\rho = \sigma_{12}\big/(\sigma_1\sigma_2)$$

حيث يتم التوصل إلى σ_{12}، وهو التغاير بين المتغيرين X_1, X_2، على النحو المبين أدناه:

$$\sigma_{12} = E(x_1 x_2) - E(x_1)E(x_2)$$

١٩-٥ التوزيع الشرطي

تحدث علاقة مهمة عندما تكون قيمة أحد المتغيرين معلومة، حيث إن x_{10} هي قيمة معيّنة للمتغير X_1 . ومع هذه الحالة الشرطية، يصبح توزيع المتغير X_2 و المُشار إليه بالرمز $x_2 \mid x_{10}$ توزيعاً طبيعياً على الشكل:

$$x_2|x_{10} \sim N\left(\mu_{x_2|x_{10}}, \sigma^2_{x_2|x_{10}}\right)$$

ويكون المتوسط والتباين الشرطي للتوزيع $x_2 \big| x_{10}$ على النحو المبين أدناه:

$$\begin{aligned} \mu_{x_2|x_{10}} &= \mu_2 + \rho(\sigma_2/\sigma_1)(x_{10} - \mu_1) \\ \sigma^2_{x_2|x_{10}} &= \sigma_2^2(1 - \rho^2) \end{aligned}$$

١٩-٦ التوزيع الطبيعي القياسي الثنائي

إنَّ أسهل طريقة لدراسة التوزيع الطبيعي الثنائي تكون من خلال التوزيع الطبيعي القياسي الثنائي ذي المتغيرين Z_1, Z_2. إنَّ التوزيعات الهامشية للمتغيرين القياسيين المتوسط صفر والتباين واحد، في حين أن الارتباط بينهما ρ يمثل مقياس العلاقة المشتركة بين المتغيرين. وتعرف الدالة بين المتغيرين على النحو التالي:

$$\left(Z_1, Z_2\right) \sim BVN(0,0,1,1,\rho)$$

مع ملاحظة أن المتوسط والانحراف المعياري لكل متغير يعرف على النحو المبين أدناه:

$$\mu_1 = 0,\ \sigma_1 = 1,\ \mu_2 = 0,\ \sigma_2 = 1$$

١٩-٧ التوزيع الهامشي

إنَّ التوزيعات الهامشية للمتغيرين تمثل توزيعاً طبيعياً قياسياً بالشكل أدناه:

$$\begin{aligned} Z_1 &\sim N(0,1) \\ Z_2 &\sim N(0,1) \end{aligned}$$

١٩-٨ التوزيعات الشرطية

عندما تكون إحدى قيم المعلمة معلومة، على سبيل المثال لتكن k_1 قيمة معيّنة للمتغير Z_1، يصبح الترميز الشرطي للمتغير Z_2 على الشكل $z_2 \mid k_1$ حيث يتم توزيع التوزيع الشرطي توزيعاً طبيعياً، ونكتب ذلك على النحو المبين أدناه:

$$z_2|k_1 \sim N(\mu_{z_2|k_1}, \sigma^2_{z_2|k_1})$$

حيث يكون المتوسط والتباين على النحو التالي:

$$\mu_{z_2|k_1} = \rho k_1$$
$$\sigma^2_{z_2|k_1} = \left(1-\rho^2\right)$$

نلاحظ أنه عندما يكون $z_2 = k_2$، يصبح المتغير الشرطي $z_1 = z_1 \Big| k_2$، والتوزيع الشرطي كذلك توزيعاً طبيعياً معرفاً على النحو المبين أدناه:

$$z_1|k_2 \sim N(\mu_{z_1|k_2}, \sigma^2_{z_1|k_2})$$

ويكون المتوسط والتباين على النحو التالي:

$$\mu_{z_1|k_2} = \rho k_2$$
$$\sigma^2_{z_1|k_2} = \left(1-\rho^2\right)$$

١٩-٩ التقريب إلى الاحتمال المشترك التراكمي

يُعبّر عن الاحتمال المشترك التراكمي للمتغير Z_1 الذي هو أقل من أو يساوي k_1، والمتغير Z_2 الذي هو أقل من أو يساوي k_2، على النحو التالي:

$$F(k_1,k_2) = P(z_1 \le k_1 \cap z_2 \le k_2)$$

وباعتبار أنه لا توجد صيغة محددة للاحتمال أعلاه، فقد تم وضع تقريب لذلك، ويتم التوصل رياضياً إلى الاحتمال المشترك على النحو التالي:

$$F(k_1,k_2) = \int_{-\infty}^{k_1}\int_{-\infty}^{k_2} f(z_1,z_2)\,dz_2 dz_1$$

وفيما يلي علاقة مكافئة أخرى للاحتمال المشترك:

$$F(k_1,k_2) = \int_{-\infty}^{k_1}\left[\int_{-\infty}^{k_2} f(z_2\middle|z_1)\,dz_2\right]f(z_1)\,dz_1$$
$$= \int_{-\infty}^{k_1} F(k_2\middle|z_1)f(z_1)\,dz_1$$

كما يستخدم التقريب العلاقتين التاليتين:

١. تقريب هاستينج الطبيعي القياسي للحصول على الاحتمال التراكمي $F(z)$ من المتغير Z كما هو موضّح في الفصل الثامن.

٢. التوزيع الطبيعي المنفصل حيث إن:

$$\mathrm{P(k) = f(k)}\Big/\sum_{\mathrm{z}=-3.0}^{3.0}\mathrm{f(z)} \quad , \quad \mathrm{z} = \left[-3.0,(0.1),3.0\right]$$

بحيث أن $f(z)$ هي دالة الكثافة الاحتمالية للمتغير $Z \sim N(0,1)$ و k هي قيمة معيّنة للمتغير Z.

يتم استخدام العلاقتين أعلاه في الحسابات التالية:

$$F(k_1,k_2) \approx \sum_{z_1=-3.0}^{k_1-0.1} F(k_2 \mid z_1)P(z_1) + 0.5\,F(k_2 \mid k_1)P(k_1)$$

على سبيل المثال، إذا كان $(k_1,k_2) = (1.0,1.0)$ فإن

$$F(1.0,1.0) = F\left(1.0\middle|-3.0\right)P\left(-3.0\right) + F\left(1.0\middle|-2.9\right)P\left(-2.9\right) + \cdots$$
$$\cdots + F\left(1.0\middle|0.9\right)P\left(0.9\right) + 1/2F\left(1.0\middle|1.0\right)P\left(1.0\right)$$

مع ملاحظة أن الدوال التراكمية موزّعة بالتوزيع الطبيعي القياسي حيث إن:

$$F\left(k_2\middle|z_1\right) = F\left[\left(k_2 - \mu_{z_2|z_1}\right)\Big/\sigma_{z_2|z_1}\right]$$

و

$$\mu_{z_2|z_1} = \rho z_1$$
$$\sigma_{z_2|z_1} = \sqrt{\left(1-\rho^2\right)}$$

جدول (١٩-١). $F(k_1,k_2)$ **عندما يكون** $\rho = -0.8$ **و** $\rho = 0.8$ **وعندما يكون** $k_1 = [-3.0,(0.5),3.0]$ **و** $k_2 = [-3.0,(0.5),3.0]$.

$\rho = -0.8$													
k_2/k_1	**-3**	**-2.5**	**-2**	**-1.5**	**-1**	**-0.5**	**0**	**0.5**	**1**	**1.5**	**2**	**2.5**	**3**
3	0	0.01	0.02	0.07	0.16	0.31	0.50	0.69	0.84	0.93	0.98	0.99	1.00
2.5	0	0.01	0.02	0.06	0.15	0.30	0.49	0.69	0.84	0.93	0.97	0.99	0.99
2	0	0	0.01	0.05	0.14	0.29	0.48	0.67	0.82	0.91	0.96	0.97	0.98
1.5	0	0	0.01	0.03	0.11	0.25	0.43	0.63	0.78	0.87	0.91	0.93	0.93
1	0	0	0	0.01	0.06	0.17	0.35	0.53	0.68	0.78	0.82	0.84	0.84
0.5	0	0	0	0	0.02	0.09	0.22	0.39	0.53	0.63	0.67	0.69	0.69
0	0	0	0	0	0.01	0.03	0.10	0.22	0.35	0.43	0.48	0.49	0.50
-0.5	0	0	0	0	0.00	0.01	0.03	0.09	0.17	0.25	0.29	0.30	0.31
-1	0	0	0	0	0.00	0.00	0.01	0.02	0.06	0.11	0.14	0.15	0.16
-1.5	0	0	0	0	0	0	0	0	0.01	0.03	0.05	0.06	0.07
-2	0	0	0	0	0	0	0	0	0	0.01	0.01	0.02	0.02
-2.5	0	0	0	0	0	0	0	0	0	0	0	0	0
-3	0	0	0	0	0	0	0	0	0	0	0	0	0
$\rho = 0.8$													
k_2/k_1	**-3**	**-2.5**	**-2**	**-1.5**	**-1**	**-0.5**	**0**	**0.5**	**1**	**1.5**	**2**	**2.5**	**3**
3	0	0.01	0.02	0.07	0.16	0.31	0.50	0.69	0.84	0.93	0.98	0.99	1.00
2.5	0	0.01	0.02	0.07	0.16	0.31	0.50	0.69	0.84	0.93	0.98	0.99	0.99
2	0	0.01	0.02	0.07	0.16	0.31	0.50	0.69	0.84	0.93	0.97	0.98	0.98
1.5	0	0.01	0.02	0.07	0.16	0.31	0.50	0.69	0.83	0.90	0.93	0.93	0.93
1	0	0.01	0.02	0.07	0.16	0.31	0.49	0.67	0.78	0.83	0.84	0.84	0.84
0.5	0	0.01	0.02	0.07	0.16	0.30	0.47	0.60	0.67	0.69	0.69	0.69	0.69
0	0	0.01	0.02	0.07	0.15	0.28	0.40	0.47	0.49	0.50	0.50	0.50	0.50
-0.5	0	0.01	0.02	0.06	0.13	0.22	0.28	0.30	0.31	0.31	0.31	0.31	0.31
-1	0	0.01	0.02	0.05	0.10	0.13	0.15	0.16	0.16	0.16	0.16	0.16	0.16
-1.5	0	0	0.02	0.03	0.05	0.06	0.07	0.07	0.07	0.07	0.07	0.07	0.07
-2	0	0	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
-2.5	0	0	0	0	0	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
-3	0	0	0	0	0	0	0	0	0	0	0	0	0

جدول (١٩-٢). $F(k_1, k_2)$ عندما يكون $\rho = (-1, 1)$ وعندما يكون $k_1 = [-3.0, (1.0), 3.0]$ و $k_2 = [-3.0, (1.0), 3.0]$.

k_1/k_2		-1.0	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0.0
3	3	1.000	0.999	0.999	0.999	0.999	0.998	0.998	0.998	0.998	0.998	0.998
3	2	0.978	0.977	0.977	0.977	0.977	0.977	0.977	0.977	0.977	0.977	0.977
3	1	0.842	0.841	0.841	0.841	0.841	0.841	0.841	0.841	0.841	0.841	0.841
3	0	0.500	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499
3	-1	0.158	0.157	0.157	0.157	0.157	0.157	0.157	0.157	0.158	0.158	0.158
3	-2	0.022	0.021	0.021	0.021	0.021	0.021	0.021	0.022	0.022	0.022	0.022
3	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	3	0.978	0.978	0.978	0.978	0.978	0.977	0.977	0.977	0.977	0.977	0.977
2	2	0.956	0.957	0.956	0.956	0.956	0.956	0.956	0.956	0.956	0.956	0.956
2	1	0.820	0.820	0.820	0.820	0.820	0.820	0.820	0.820	.0821	0.822	0.823
2	0	0.478	0.478	0.478	0.478	0.479	0.480	0.481	0.483	0.484	0.487	0.489
2	-1	0.136	0.137	0.138	0.140	0.143	0.145	0.147	0.150	0.151	0.153	0.154
2	-2	0.001	0.009	0.013	0.015	0.017	0.018	0.019	0.020	0.020	0.021	0.021
2	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	3	0.842	0.842	0.842	0.842	0.842	0.842	0.841	0.841	0.841	0.841	0.841
1	2	0.820	0.820	0.820	0.820	0.820	0.820	0.820	0.821	0.821	0.822	0.823
1	1	0.684	0.684	0.684	0.685	0.686	0.688	0.690	0.694	0.698	0.703	0.708
1	0	0.342	0.343	0.348	0.355	0.364	0.373	0.383	0.392	0.402	0.411	0.421
1	-1	0.006	0.044	0.061	0.075	0.086	0.096	0.105	0.113	0.120	0.127	0.133
1	-2	0.000	0.000	0.002	0.004	0.007	0.009	0.012	0.014	0.015	0.017	0.018
1	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0	3	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
0	2	0.478	0.478	0.478	0.479	0.479	0.480	0.482	0.483	0.485	0.487	0.489
0	1	0.342	0.343	0.348	0.355	0.364	0.373	0.383	0.392	0.402	0.411	0.421
0	0	0.010	0.072	0.103	0.127	0.148	0.167	0.185	0.202	0.218	0.234	0.250
0	-1	0.000	0.001	0.006	0.013	0.022	0.031	0.041	0.050	0.060	0.070	0.079
0	-2	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.005	0.007	0.009	0.011
0	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-1	3	0.158	0.158	0.158	0.158	0.158	0.158	0.158	0.158	0.158	0.159	0.159
-1	2	0.136	0.136	0.138	0.140	0.143	0.146	0.148	0.150	0.152	0.154	0.155
-1	1	0.006	0.043	0.061	0.075	0.086	0.096	0.105	0.113	0.121	0.127	0.134

تابع جدول (١٩-٢).

k_1/k_2		-1.0	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0.0
-1	0	0.000	0.001	0.006	0.013	0.022	0.031	0.041	0.050	0.060	0.070	0.079
-1	-1	0.000	0.000	0.000	0.001	0.002	0.004	0.007	0.010	0.015	0.020	0.025
-1	-2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.003
-1	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-2	3	0.022	0.021	0.022	0.022	0.022	0.022	0.022	0.023	0.023	0.023	0.023
-2	2	0.001	0.009	0.013	0.015	0.017	0.019	0.020	0.021	0.021	0.022	0.022
-2	1	0.000	0.000	0.002	0.004	0.007	0.010	0.012	0.014	0.016	0.018	0.019
-2	0	0.000	0.000	0.000	0.000	0.001	0.002	0.003	0.005	0.007	0.009	0.011
-2	-1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.002	0.004
-2	-2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-2	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-3	3	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
-3	2	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001
-3	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001
-3	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001
-3	-1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-3	-2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-3	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

k_1/k_2		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
3	3	0.998	0.998	0.998	0.998	0.998	0.999	0.999	0.999	0.999	0.999	1.000
3	2	0.977	0.977	0.977	0.977	0.977	0.977	0.978	0.978	0.978	0.978	0.978
3	1	0.841	0.841	0.841	0.841	0.841	0.842	0.842	0.842	0.842	0.842	0.842
3	0	0.499	0.499	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
3	-1	0.158	0.158	0.158	0.158	0.158	0.158	0.158	0.158	0.158	0.158	0.158
3	-2	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022
3	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	3	0.977	0.977	0.977	0.977	0.977	0.977	0.978	0.978	0.978	0.978	0.978
2	2	0.956	0.956	0.957	0.957	0.958	0.960	0.961	0.963	0.965	0.969	0.977
2	1	0.823	0.824	0.826	0.828	0.830	0.832	0.835	0.838	0.840	0.842	0.842
2	0	0.489	0.491	0.493	0.495	0.497	0.498	0.499	0.500	0.500	0.500	0.500
2	-1	0.154	0.156	0.157	0.157	0.158	0.158	0.158	0.158	0.158	0.158	0.158
2	-2	0.021	0.021	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022

تابع جدول (١٩-٢).

k_1/k_2		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
2	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	3	0.841	0.841	0.841	0.841	0.842	0.842	0.842	0.842	0.842	0.842	0.842
1	2	0.823	0.824	0.826	0.828	0.830	0.833	0.835	0.838	0.840	0.842	0.842
1	1	0.708	0.714	0.721	0.729	0.737	0.746	0.756	0.767	0.781	0.799	0.836
1	0	0.421	0.430	0.440	0.450	0.459	0.469	0.478	0.487	0.494	0.499	0.500
1	-1	0.133	0.138	0.143	0.148	0.151	0.154	0.156	0.158	0.158	0.158	0.158
1	-2	0.018	0.019	0.020	0.021	0.021	0.022	0.022	0.022	0.022	0.022	0.022
1	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0	3	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
0	2	0.489	0.491	0.493	0.495	0.497	0.498	0.499	0.500	0.500	0.500	0.500
0	1	0.421	0.430	0.440	0.450	0.459	0.469	0.478	0.487	0.494	0.499	0.500
0	0	0.250	0.266	0.282	0.298	0.315	0.333	0.352	0.373	0.397	0.428	0.490
0	-1	0.079	0.089	0.098	0.108	0.117	0.127	0.136	0.145	0.152	0.157	0.158
0	-2	0.011	0.013	0.015	0.017	0.018	0.020	0.021	0.021	0.022	0.022	0.022
0	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-1	3	0.159	0.159	0.159	0.158	0.158	0.158	0.158	0.158	0.158	0.158	0.158
-1	2	0.155	0.156	0.157	0.158	0.158	0.158	0.158	0.158	0.158	0.158	0.158
-1	1	0.134	0.139	0.144	0.148	0.152	0.154	0.156	0.157	0.158	0.158	0.158
-1	0	0.079	0.089	0.098	0.108	0.117	0.127	0.136	0.145	0.152	0.157	0.158
-1	-1	0.025	0.031	0.038	0.045	0.053	0.062	0.072	0.083	0.097	0.115	0.152
-1	-2	0.003	0.005	0.006	0008	0.010	0.012	0.015	0.017	0.020	0.021	0.022
-1	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-2	3	0.023	0.023	0.023	0.023	0.023	0.022	0.022	0.022	0.022	0.022	0.022
-2	2	0.022	0.022	0.023	0.023	0.023	0.022	0.022	0.022	0.022	0.022	0.022
-2	1	0.019	0.020	0.021	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022
-2	0	0.011	0.013	0.016	0.017	0.019	0.020	0.021	0.022	0.022	0.022	0.022
-2	-1	0.004	0.005	0.007	0.009	0.011	0.013	0.015	0.018	0.020	0.021	0.022
-2	-2	0.000	0.001	0.001	0.002	0.003	0.004	0.005	0.007	0.009	0.012	0.020
-2	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-3	3	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000
-3	2	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000
-3	1	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000

تابع جدول (١٩-٢).

k_1/k_2		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
-3	0	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000
-3	-1	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000
-3	-2	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.000
-3	-3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

١٩-١٠ الجداول الإحصائية

يبيّن الجدول (١٩-١) الاحتمال التراكمي $F(k_1,k_2)$ بالنسبة للقيم المختارة لـ k_1,k_2، وعندما يكون $\rho = -0.8$ و $\rho = 0.8$. يبيّن الجدول (١٩-٢) الاحتمال التراكمي $F(k_1,k_2)$ بالنسبة للقيم المختارة لـ k_1,k_2 وبالنسبة لـقيمة ρ التي تتراوح ما بين (1+ ,1−).

المثال (١٩-١). يوجد لدى المحلل البيانات الثنائية $\left(X_1,X_2\right) \sim BVN(5,10,1,2,0.8)$ ويريد إيجاد احتمال X_1 الذي هو أقل من 6 أو يساويها، و X_2 الذي هو أقل من 13 أو يساويها. للتوضيح والتبسيط، يكون الاحتمال التراكمي للمتغيرين X_1,X_2 على النحو التالي: $F_x = (6.13)$. ويصبح الاحتمال التراكمي للمتغيرين المقابلين Z_1,Z_2 على الشكل $F_z(1.0,1.5)$، حيث إن:

$$\begin{aligned} z_1 &= (6.0-5.0)/1.0 = 1.0 \\ z_2 &= (13.0-10.0)/2.0 = 1.5 \end{aligned}$$

باستخدام الجدول (١٩-١)، يصبح الاحتمال المطلوب:

$$F_x(6.13) = F_z(1.0,1.5) = 0.83$$

المثال (١٩-٢). عند تطبيق التوزيع الطبيعي الثنائي المبين في المثال (١٩-١)، وبافتراض أن المحلل يسعى إلى إيجاد احتمال X_1 بين 5 و6، و X_2 بين 10 و13. باستخدام الجدول (١٩-١) وذا الترميز المبين في المثال السابق، يتم التوصل إلى الاحتمال كما هو مبين أدناه:

$$P(5 \le X_1 \le 6 \cap 10 \le X_2 \le 13)$$

$$
\begin{aligned}
&= F_x(6,13) - F_x(6,10) - F_x(5,13) + F_x(5,10) \\
&= F_x(1.0,1.5) - F_x(1.0,0.0) - F_x(0.0,1.5) + F_x(0.0,0.0) \\
&= 0.83 - 0.49 - 0.50 + 0.40 \\
&= 0.24
\end{aligned}
$$

١٩-١١ ملخص

يحتوي التوزيع الطبيعي الثنائي على المتغيرين X_1, X_2، المرتبطين بصورة مشتركة. هذا التوزيع معرّف بخمسة معالم حيث تمَّ توضيح التوزيعين الهامشي والشرطي مع متوسط وتباين كل توزيع. كذلك تم التحويل إلى المتغيرين Z_1, Z_2، من المتغيرين X_1, X_2، حيث يسهل تطبيقهما في الحسابات كما أنَّهما مرتبطان من خلال التوزيع الطبيعي القياسي الثنائي، حيث تم توضيح التوزيعات الهامشية والشرطية المقابلة. بعد ذلك تم وضع طريقة تقريبية لحساب الاحتمال التراكمي المشترك للمتغيرين وعرض بعض الجداول. كذلك تم عرض بعض الأمثلة التوضيحية.

الفصل العشرون

توزيع اللوغاريتم الطبيعي الثنائي
BIVARIATE LOGNORMAL DISTRIBUTION

٢٠-١ مقدمة

يبيّن المؤلف في كتاب ثوموبولوس ولونغيناو عام (1984)، [Thomopoulos and Longinow (1984). p 3045–3049] كيفية حساب الاحتمال التراكمي لتوزيع اللوغاريتم الطبيعي الثنائي في مسألة موثوقية الهيكلة الهندسية. ويبدو توزيع اللوغاريتم الطبيعي الثنائي ذي المتغيرين X_1, X_2 للوهلة الأولى من الصعب استخدامه، إلا أنه عند أخذ اللوغاريتم الطبيعي لكل متغير، يظهر التوزيع الطبيعي الثنائي وهو سهل الاستخدام، وتصبح المعالم الخمسة للتوزيع الطبيعي الثنائي هي معالم توزيع اللوغاريتم الطبيعي الثنائي. يبيّن الفصل كيفية تحويل المعالم من اللوغاريتم الطبيعي الثنائي إلى الطبيعي الثنائي والعكس. كما يبيّن كيفية حساب الترابط بين المتغيرين في اللوغاريتم الطبيعي الثنائي، وكيفية حساب الاحتمال المشترك لأي زوج (x_1, x_2). وأخيراً تم تقديم مثال لمساعدة القارئ في المنهجية المستخدمة.

٢٠-٢ أساسيات

عندما يرتبط المتغيران X_1, X_2 بصورةٍ مشتركة من خلال توزيع اللوغاريتم الطبيعي الثنائي، فإنه يكون لهما خمسة معالم هي $(\mu_{y1}, \mu_{y2}, \sigma_{y1}, \sigma_{y2}, \rho_y)$. ونكتب اختصاراً:

$$\left(X_1, X_2\right) \sim BVLN(\mu_{y1}, \mu_{y2}, \sigma_{y1}, \sigma_{y2}, \sigma_y)$$

إنَّ التوزيعات الهامشية للمتغيرين تكون لوغاريتماً طبيعياً، وتكتب على النحو المبين أدناه:

$$X_1 \sim LN(\mu_{y1}, \sigma_{y1})$$
$$X_2 \sim LN(\mu_{y2}, \sigma_{y2})$$

كما أنَّ معلمتي المتوسط والتباين أعلاه مأخوذتان من التوزيعات الطبيعية المقابلة لكل من $y_1 = \ln(x_1)$ و $y_2 = \ln(x_2)$ حيث إن $\ln$ هو اللوغاريتم الطبيعي وتكتب على النحو التالي:

$$Y_1 \sim N(\mu_{y1}, \sigma_{y1}^2)$$
$$Y_2 \sim N(\mu_{y2}, \sigma_{y2}^2)$$

إنَّ العلاقة بين Y_1, Y_2 هي ρ_y. كما أن المتغيرين هما متغيرا التوزيع الطبيعي الثنائي وتكتب على النحو التالي:

$$\left(Y_1, Y_2\right) \sim BVN(\mu_{y1}, \mu_{y2}, \sigma_{y1}, \sigma_{y2}, \rho_y)$$

ويمكن تحويل متوسط وتباين X من متوسط وتباين Y كما هو مبين أدناه:

متوسط المتغير x = $\mu_x = \exp\left[\mu_y + \sigma_y^2/2\right]$

تباين المتغير x = $\sigma_x^2 = \exp\left[2\mu_y + \sigma_y^2\right]\left[\exp(\sigma_y^2) - 1\right]$

وبالمثل يمكن التوصل إلى متوسط وتباين متغير Y من متوسط وتباين المتغير X على النحو التالي:

متوسط المتغير y = $\mu_y = \ln\left[\mu_x^2 \Big/ \sqrt{\mu_x^2 + \sigma_x^2}\right]$

تباين المتغير y = $\sigma_y^2 = \ln\left[1 + \sigma_x^2 / \mu_x^2\right]$

كما تم إدراج التغاير والترابط بين المتغيرين الطبيعيين Y_1, Y_2، أدناه:

$$\sigma_{y1y2} = E(y_1 y_2) - E(y_1)E(y_2)$$
$$p_{y1y2} = \sigma_{y1y2} \Big/ \sigma_{y1}\sigma_{y2}$$

وتكون العلاقة بين المتغيرين اللوغاريتميين الطبيعيين X_1, X_2 على النحو التالي:

$$\rho_{x1x2} = \left[\exp(\sigma_{y1y2}) - 1 \;/\; \left\{\left[\exp(\sigma_{y1}^2 - 1)\right]\left[\exp(\sigma_{y2}^2 - 1)\right]\right\}^{0.5}\right.$$

٢٠-٣ الاحتمال التراكمي

لتوضيح دالة الاحتمال التراكمي، تم ترميزها على النحو التالي:

$F_x(x_1, x_2)$ هو الاحتمال التراكمي لمتغيرات اللوغاريتم الطبيعي X.

$F_y(y_1, y_2)$ هو الاحتمال التراكمي لمتغيرات التوزيع الطبيعي Y.

$F_z(z_1, z_2)$ هو الاحتمال التراكمي لمتغيرات التوزيع الطبيعي القياسي Z.

ويُشار إلى دالة الاحتمال التراكمي للقيمة x_{10} بالنسبة للمتغير X_1، والقيمة x_{20} بالنسبة للمتغير X_2، كما يلي:

$$F_x(x_{10}, x_{20}) = P\left(x_1 \le x_{10} \cap x_2 \le x_{20}\right)$$

وللحصول على الاحتمال أعلاه، يتم اتباع الخطوات الأربع التالية:

١. تحويل قيم اللوغاريتم الطبيعي للمتغير X إلى قيم التوزيع الطبيعي للمتغير Y من خلال العلاقات التالية:

$$\begin{aligned} y_{10} &= \ln(x_{10}) \\ y_{20} &= \ln(x_{20}) \end{aligned}$$

٢. تغيير قيم التوزيع الطبيعي للمتغير Y إلى قيم التوزيع الطبيعي القياسي للمتغير k كما يلي:

$$\begin{aligned} k_1 &= \left[y_{10} - \mu_{y1}\right] / \sigma_{y1} \\ k_2 &= \left[y_{20} - \mu_{y2}\right] / \sigma_{y2} \end{aligned}$$

٣. باستخدام التوزيع الطبيعي القياسي الثنائي، ومن خلال ρ_y، يتم حساب $F_z(k_1, k_2)$ على النحو المبين في الفصل التاسع عشر.

٤. وأخيراً

$$F_x(x_{10}, x_{20}) = F_z(k_1, k_2)$$

المثال (٢٠-١). يوجد لدى المحلل بيانات توزيع اللوغاريتم الطبيعي ثنائي المتغيرات التي تتضمن $(X_1, X_2) \sim BVLN(2, 4, 1, 2, 0.8)$. في هذه الحالة، يبيّن الجدول (٢٠-١) الاحتمالات

التراكمية مع القيم المختارة k_1, k_2 و x_1, x_2 عندما يكون $\rho_y = 0.8$. تم أخذ الاحتمالات الجدولية من الجدول (١٩-١) في الفصل التاسع عشر. يتم التوصل إلى قيم x_1, x_2 على النحو التالي:

$$x_1 = \exp(2.0 + k_1 \times 1.0)$$
$$x_2 = \exp(4.0 + k_2 \times 2.0)$$

جدول (٢٠-١). $F_x(x_1, x_2)$ **بالنسبة للقيم المختارة لـ** $(x_1, x_2) \sim BVLN(2,4,1,2,0.8)$ **و** $F_z(k_1, k_2)$ **النسبة لقيم** $(k_1, k_2) \sim BVN(0,0,1,1,0.8)$.

$\rho = 0.8$														
k_2/k_1	**-3**	**-2.5**	**-2**	**-1.5**	**-1**	**-0.5**	**0**	**0.5**	**1**	**1.5**	**2**	**2.5**	**3**	**×2**
3	0	0.01	0.02	0.07	0.16	0.31	0.50	0.69	0.84	0.93	0.98	0.99	1.00	22.026
2.5	0	0.01	0.02	0.07	0.16	0.31	0.50	0.69	0.84	0.93	0.98	0.99	0.99	8103
2	0	0.01	0.02	0.07	0.16	0.31	0.50	0.69	0.84	0.93	0.96	0.98	0.98	2980
1.5	0	0.01	0.02	0.07	0.16	0.31	0.50	0.69	0.83	0.90	0.93	0.93	0.93	1096
1	0	0.01	0.02	0.07	0.16	0.31	0.49	0.67	0.78	0.83	0.84	0.84	0.84	403
0.5	0	0.01	0.02	0.07	0.16	0.30	0.47	0.60	0.67	0.69	0.69	0.69	0.69	148
0	0	0.01	0.02	0.07	0.15	0.28	0.40	0.47	0.49	0.50	0.50	0.50	0.50	55
-0.5	0	0.01	0.02	0.06	0.13	0.22	0.28	0.30	0.31	0.31	0.31	0.31	0.31	20
-1	0	0.01	0.02	0.05	0.10	0.13	0.15	0.16	0.16	0.16	0.16	0.16	0.16	7
-1.5	0	0	0.02	0.03	0.05	0.06	.07	0.07	0.07	0.07	0.07	0.07	0.07	3
-2	0	0	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	1
-2.5	0	0	0	0	0	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.4
-3	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1
×1	0.4	0.6	1	2	3	4	7	12	20	33	55	90	148	

على افتراض أنَّ المحلل يريد إيجاد الاحتمال التراكمي المشترك عندما يكون $x_1 = 30,\ x_2 = 500$، أي $F_x(30, 500)$. لتحقيق ذلك، يتم اتخاذ الخطوات الأربع التالية:

١. تحويل قيم اللوغاريتم الطبيعي x إلى قيم التوزيع الطبيعي كما هو مبين أدناه:

$$y_{10} = \ln(30) = 3.40$$
$$y_{20} = \ln(500) = 6.21$$

٢. تحويل قيم التوزيع الطبيعي إلى قيم التوزيع الطبيعي القياسي على النحو التالي:

$$k_1 = (3.40 - 2.00)/1.00 = 1.40$$
$$k_2 = (6.21 - 4.00)/2.00 = 1.10$$

٣. باستخدام الجدول (٢٠-١) يتم إيجاد الاحتمال التراكمي المطلوب على النحو المبين أدناه:

$$F_z(1.40, 1.1) \approx 0.82$$

٤. وأخيراً

$$F_x(30, 500) = F_y(3.40, 6.21) = F_z(1.40, 1.1) \approx 0.82$$

٢٠-٤ ملخص

يحتوي توزيع اللوغاريتم الطبيعي الثنائي على متغيرين هما X_1, X_2، وتعد التوزيعات الهامشية توزيعات لوغاريتم طبيعي. وعند تحويل المتغيرين من خلال أخذ اللوغاريتم الطبيعي لكل منهما، يظهر التوزيع الطبيعي الثنائي. كما تصبح المعالم الخمسة للتوزيع الأخير معالم لتوزيع اللوغاريتم الطبيعي الثنائي. وتتمثل طريقة اشتقاق الاحتمال المشترك لزوج متغيري توزيع اللوغاريتم الطبيعي الثنائي في حساب الاحتمال المشترك المقابل لمتغيري التوزيع الطبيعي الثنائي.

المراجع

Abramowitz, M. & Stegun, I. (1964), Handbook of Mathematical Functions, National Bureau of Standards, Washington, D.C.

Aitchison, J. & Brown, J. (1969). The Lognormal Distribution. Cambridge: Cambridge University Press.

Beyer, W.H. (1968). Handbook of Tables for Probability and Statistics. Chemical Rubber Co.

Broadbent, S. (1956). Lognormal Approximation to Products and Quotients. 43, 404-417: Biometrika.

Brown, R.G. (1959). Smoothing, Forecasting and Prediction of Discrete Time Series. Engelwood Cliffs, N.J.: Prentice Hall.

Crow, E. & Shinizu, K. (1988). Lognormal Distribution: Theory and Applications, New York: Marcel Dekker.

Galton, F. (1909). Memories of My Life. New York: E.P. Dutton & Co.

Hasting, N.A.J. & Peacock, J.B. (1974). Statistical Distributions. New York: Wiley & Sons

Hines, W.W., Montgomery, L.D.C., Goldsman, D.M., & Burror, C.M. (2003). Probability and Statistics for Engineers. New York: Wiley & Sons.

Jantaravareerat, M. (1998). Approximation of the Distribution Function for the Standard Bivariate Normal. Doctoral Dissertation, Stuart School of Business: Illinois Institute of Technology.

Jantaravareerat, M. & Thomopoulos N. (1998). Some New Tables on the Standard Normal Distribution. 30, pp179-184: Computing Science and Statistics.

Johnson, A.C. (2001). On The Truncated Normal Distribution. Doctoral Dissertation. Stuart School of Business: Illinois Institute of Technology.

Johnson A.C., & Thomopoulos, N.T. (2002). Characteristics and Tables of the Left-truncated Normal Distribution. Proceedings of the Midwest Decision Sciences Institute, 133-139.

Law, A. & Kelton, W. (2000). Simulation, Modeling and Analysis. Boston: McGraw Hill.

Lindee, C., (2001). The Multivariate Standard Normal Distribution. Doctoral Dissertation. Stuart School of Business: Illinois Institute of Technology

Lindee, C. & Thomopoulos, N. (2001). Values for the Cumulative Distribution Function of the Standard Multivariate Normal Distribution. 48, pp 600-609: Proceedings of the Midwest Decision Sciences Institute.

Robson, D.S., & Whitlock, J.W. (1964). Estimation of a Truncation Point, 51, pp 33-39: Biometrical.

Schneider, J. (1986). Truncated and Censored Samples from Normal Populations. New York: Marcel Dekker.

Thomopoulos, N.T. (1980). Applied Forecasting Methods. Englewood Cliffs, N.J.: Prentice Hall.

Thomopoulos, N.T. (2013). Essentials of Monte Carlo Simulation. New York: Springer.

Thomopoulos, N.T. (2016). Demand Forecasting for Inventory Control. New York: Springer.

Thomopoulos, N.T., & Longinow, A.C. (1984). Bivariate Lognormal Probability Distribution. 110: Journal of Structural Engineering.

Thomopoulos, N.T. & Johnson, A.C. (2003). Tables and Characteristics of the Standardized Lognormal Distribution. 103, pp 1-6: Proceeding of the Decision Sciences Institute

Thomopoulos, N.T., & Johnson, A.C. (2004). Some Measures on the Standard Bivariate Lognormal Distribution; pp 1721-1726: Proceedings of the Decision Sciences Institute.

Zanakis, S.H. (1979). A simulation study of some simple estimators for the three parameter Weibull distribution. J. Stat Comput. Simul.

ثبت المصطلحات

أولًا: عربي - إنجليزي

أ

الاحتمال الشرطي	Conditional probability
الاحتمال المكمل	Complementary probability
احتمال متساوٍ	Equal probability
الارتباط الأسي	Exponential connection
أساسيات	Fundamentals
إخفاقات	Fails
انحدار	Regression
إرجاع	Replacement
الانحراف المعياري	Standard deviation
الالتواء نحو اليسار	Skewed left
الالتواء نحو اليمين	Skewed right

ب

بيانات العينة	Sample data

ت

توزيع بيتا	Beta distribution

Binomial distribution	توزيع ثنائي الحدين
Bivariate normal distribution	التوزيع الطبيعي الثنائي
Bivariate lognormal distribution	توزيع اللوغاريتم الطبيعي الثنائي
Continuous distribution	التوزيع المتصل
Covariance	التغاير
Conditional distribution	التوزيع الشرطي
Continuous Uniform distribution	التوزيع المتصل المنتظم
Discrete distribution	التوزيع المتقطع
Discrete Uniform distribution	التوزيع المنتظم المتقطع
Estimation	تقدير
Exponential distribution	التوزيع الأسي
Erlang distribution	توزيع إيرلانج
Galton distribution	توزيع جالتون
Gamma distribution	توزيع جاما
Geometric distribution	التوزيع الهندسي
Hastings approximations	تقريبات هاستينج
Hyper Geometric distribution	التوزيع الهندسي الزائد
Lognormal distribution	توزيع اللوغاريتم الطبيعي
Left truncated normal distribution	التوزيع الطبيعي المقطوع الأيسر
Marginal distribution	التوزيع الهامشي
Negative binomial distribution	توزيع ثنائي الحدين السالب
Normal distribution	التوزيع الطبيعي
Pascal distribution	توزيع باسكال
Poisson distribution	توزيع بواسون
Right truncated normal distribution	التوزيع الطبيعي المقطوع الأيمن

التوزيع الإحصائي	Statistical distribution
التوزيع الطبيعي القياسي	Standard normal distribution
التوزيع المثلثي	Standard triangular distribution
تحويل المتغيرات	Transforming variables
التوزيع المثلثي القياسي	Triangular distribution
التباين	Variance
توزيع ويبل	Weibull distribution

ج

الجبر	Algebra
جداول	Tables

ح

الحد الأدنى	Lower limit
الحد الأعلى	Upper limit

خ

خوارزمية	Algorithm
خبير	Expert
خاصية فقدان الذاكرة	Memory less property

د

دالة التوزيع التراكمية	Cumulative probability distribution
دالة جاما	Gamma function
دالة الكثافة الاحتمالية	Probability density function
دالة الكتلة الاحتمالية	Probability mass function

ز

زوجي	Even

زمن الإخفاق	Time to fail

س

سيناريو	Scenario

ش

شكل الجرس	Bell-shaped

ص

صفوف الانتظار	Queuing theory
صيغة زاناكس	Zanakis formula

ط

طريقة المربعات الصغرى	Method of least squares
طريقة الإمكان الأعظم	Maximum likelihood method

ع

العدد الأعلى	Ceiling
العلاقات المكافئة	Equivalent relations
العدد الأدنى	Floor
عدد صحيح	Integer
العزوم	Moments
عدد المحاولات	Number of trials
علاقة بواسون	Poisson relation
علاقة	Relation
عينة	Sample

ف

فشل	Failure
فردي	Odd

ق

القيمة المتوقعة	Expected value
قياسات	Measures
القمة	Peak
القيم الجدولية	Table values

م

معامل الاختلاف	Coefficient of variation
مراقبة المخزون	Inventory control
معلوم	Known
المتوسط	Mean
المنوال	Mode
مقدر الإمكان الأعظم	Maximum likelihood estimator
مشاهدات	Observations
المعلمة	Parameter
مجتمع	Population
المتغير العشوائي	Random variable
المدى	Range
الموثوقية	Reliability
مراجع	Reference
معلمة المقياس	Scale parameter
معلمة الشكل	Shape parameter
متماثل	Symmetrical
محاكاة	Simulation
المفاهيم الإحصائية	Statistical concepts

ملخص	Summary
مجهول	Unknown
منتظم	Uniform

ن

نظرية النهاية المركزية	Central limit theorem
نسبة لكسيس	Lexis ratio
النقطة المئوية لمتغير	Percent point of a random variable
نجاح	Success

و

الوسيط	Median

ي

يتنبأ	Predict

ثانيًا: إنجليزي – عربي

A

Algebra	الجبر
Algorithm	خوارزمية

B

Beta distribution	توزيع بيتا
Binomial distribution	توزيع ثنائي الحدين
Bivariate normal distribution	التوزيع الطبيعي الثنائي
Bivariate lognormal distribution	توزيع اللوغاريتم الطبيعي الثنائي
Bell-shaped	شكل الجرس

C

Conditional probability	الاحتمال الشرطي
Complementary probability	الاحتمال المكمل
Continuous distribution	التوزيع المتصل
Covariance	التغاير
Conditional distribution	التوزيع الشرطي
Continuous Uniform distribution	التوزيع المتصل المنتظم
Cumulative probability distribution	دالة التوزيع التراكمية
Ceiling	العدد الأعلى
Coefficient of variation	معامل الاختلاف
Central limit theorem	نظرية النهاية المركزية

D

Discrete distribution	التوزيع المتقطع

التوزيع المنتظم المتقطع	Discrete Uniform distribution

E

احتمال متساوٍ	Equal probability
الارتباط الأسي	Exponential connection
تقدير	Estimation
التوزيع الأسي	Exponential distribution
توزيع إيرلانج	Erlang distribution
خبير	Expert
زوجي	Even
العلاقات المكافئة	Equivalent relations
القيمة المتوقعة	Expected value

F

أساسيات	Fundamentals
إخفاقات	Fails
العدد الصحيح الأدنى	Floor
فشل	Failure

G

توزيع جالتون	Galton distribution
توزيع جاما	Gamma distribution
التوزيع الهندسي	Geometric distribution
دالة جاما	Gamma function

H

تقريبات هاستينج	Hastings approximations
التوزيع الهندسي الزائد	Hyper Geometric distribution

I

Integer	عدد صحيح
Inventory control	مراقبة المخزون

K

Known	معلوم

L

Lognormal distribution	توزيع اللوغاريتم الطبيعي
Left truncated normal distribution	التوزيع الطبيعي المقطوع الأيسر
Lower limit	الحد الأدنى
Lexis ratio	نسبة لكسيس

M

Marginal distribution	التوزيع الهامشي
Memory less property	خاصية فقدان الذاكرة
Method of least squares	طريقة المربعات الصغرى
Maximum likelihood method	طريقة الإمكان الأعظم
Moments	العزوم
Measures	قياسات
Mean	المتوسط
Mode	المنوال
Maximum likelihood estimator	مقدر الإمكان الأعظم
Median	الوسيط

N

Negative binomial distribution	توزيع ثنائي الحدين السالب

التوزيع الطبيعي	Normal distribution
عدد المحاولات	Number of trials

O

فردي	Odd
مشاهدات	Observations

P

توزيع باسكال	Pascal distribution
توزيع بواسون	Poisson distribution
دالة الكثافة الاحتمالية	Probability density function
دالة الكتلة الاحتمالية	Probability mass function
علاقة بواسون	Poisson relation
القمة	Peak
المعلمة	Parameter
مجتمع	Population
النقطة المئوية لمتغير	Percent point of a random variable
يتنبأ	Predict

Q

صفوف الانتظار	Queuing theory

R

انحدار	Regression
إرجاع	Replacement
التوزيع الطبيعي المقطوع الأيمن	Right truncated normal distribution
علاقة	Relation
المتغير العشوائي	Random variable

Range	المدى
Reliability	الموثوقية
Reference	مراجع

S

Standard deviation	الانحراف المعياري
Skewed left	الالتواء نحو اليسار
Skewed right	الالتواء نحو اليمين
Sample data	بيانات العينة
Statistical distribution	التوزيع الإحصائي
Standard normal distribution	التوزيع الطبيعي القياسي
Standard triangular distribution	التوزيع المثلثي القياسي
Sample	عينة
Scale parameter	معلمة المقياس
Shape parameter	معلمة الشكل
Symmetrical	متماثل
Simulation	محاكاة
Statistical concepts	المفاهيم الإحصائية
Summary	ملخص
Success	نجاح
Scenario	سيناريو

T

Transforming variables	تحويل المتغيرات
Triangular distribution	التوزيع المثلثي القياسي
Tables	جداول

زمن الإخفاق	Time to fail
القيم الجدولية	Table values

U

الحد الأعلى	Upper limit
مجهول	Unknown
منتظم	Uniform

V

التباين	Variance

W

توزيع ويبل	Weibull distribution

Z

صيغة زاناكس	Zanakis formula

كشاف الموضوعات

ح

خ

د

ز

ط

ع

ق

م

ن

و